AD-A100 948 PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS F/G 12/1
SOME RESULTS IN THE THEORY OF SUBSET SELECTION PROCEDURES.(U)
AUG 80 H YANG N00014-75-C-0455
UNCLASSIFIED MMS-80-22 NL

1 OF 2
AD A
100948

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

LEVEL II

1

Some Results in the Theory
of Subset Selection Procedures

by

Hwa-Ming Yang
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #80-22

August, 1980

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

INTRODUCTION

Selection and ranking (ordering) problems in statistical inference arise mainly because the classical tests of homogeneity are often inadequate in certain situations where the experimenter is interested in comparing k (≥ 2) populations, treatments or processes with the goal of selecting one or more worthwhile (good) populations. Mosteller (1948), Paulson (1949), Bahadur (1950) and Bahadur and Robbins (1950) were among the earliest research workers to recognize this inadequacy and to formulate the problem as a multiple decision problem aimed at the selection and ranking of the k populations.

In the thirty years since these early papers, selection and ranking problems have become an active area of statistical research. There have been two approaches to these problems, the 'indifference zone' approach and the 'subset selection' approach. In the first approach, due to Bechhofer (1954), the experimenter wishes to select one population (or a fixed number $t \geq 1$ of population) which is guaranteed to be the one of interest to him with a fixed probability P^* whenever the unknown parameters lie outside some subspace of the parameter space, the so called indifference zone. Important contributions using this approach have been made by Bechhofer and Sobel (1954), Bechhofer, Dunnett and Sobel (1954), Sobel and Huyett (1957), Sobel (1967), Bechhofer, Kiefer and Sobel (1968), Mahamunulu (1967), Desu and Sobel (1968, 1971) and

Tamhane and Bechhofer (1977, 1979) among others. A quite complete bibliography may be found in Gupta and Panchapakesan (1979) (see also Gibbons, Olkin and Sobel (1977)).

The second approach pioneered by Gupta (1956, 1963, 1965) assumes no a priori information about the parameter space. A single population is not necessarily chosen; rather a subset of the given k populations is selected depending on the outcome of the experiment. It is guaranteed to contain the population(s) of interest with probability which is at least equal to P* (the basic probability requirement) regardless of the true unknown configurations of the parameters. Some recent contributors in the category of subset selection include: Deely (1965), Gnanadesikan (1966), Gnanadesikan and Gupta (1970), Gupta (1967), Gupta and Studden (1970), Nagel (1970), Gupta and Nagel (1971), Gupta and Panchapakesan (1972), Rizvi and Sobel (1967), McDonald (1969), Gupta and McDonald (1970), Santner (1975), W. T. Huang (1972), D. Y. Huang (1975), Gupta and Huang (1975a, 1975b) and Gupta and Huang (1976).

Subset selection procedures can also be thought of as screening procedures which enable the experimenter to select a subset of populations (under study) which contains the populations of interest so that the populations in the selected subset can be further studies.

Sequential and multistage aspects of the ranking and selection problems, have been explored, based on the indifference zone approach by Bechhofer, Dunnett and Sobel (1954), Bechhofer (1958), Paulson (1962, 1963, 1964, 1967) and Bechhofer, Kiefer and Sobel (1968). Barron and Gupta (1972), Huang (1972), Gupta and Huang (1975). Gupta and Miescke (1979) and Carroll (1974) have investigated subset selection procedures, based on sequential sampling.

Contributions to optimum properties of subset selection procedures have been made by Lehmann (1961), Studden (1967), Deely and Gupta (1968), Berger (1977, 1979), Gupta and Hsu (1978), Gupta and Miescke (1978), Berger and Gupta (1980).

In the decision-theoretic approach to the subset selection problems, Goel and Rubin (1977), Chernoff and Yahav (1977), Bickel and Yahav (1977), Gupta and Hsu (1978), Miescke (1979), Gupta and Kim (1980), Gupta and Hsiao (1980) have given different formulations under different loss functions and carried out investigations which indicate that the Gupta type maximum (minimum) means procedures are quite 'optimal' and 'robust'.

The main purpose of this thesis is to study some problems using the subset selection approach and provide procedures and results for some unsolved problems.

Chapter I considers the problem of selecting a subset containing all populations better than a control under an ordering prior. Here, by an ordering prior we mean that there exists a known simple or partial order relationship among the unknown parameters of the treatmetns (excluding the control). Three new selection procedures are proposed and studied. These procedures do meet the usual requirement that the probability of a correct selection is greater than or equal to a pre-determined number P*. Two of the three procedures use the isotonic regression over the sample means of the k treatments with respect to (wrt) the given ordering prior. Tables which are necessary to carry out the selection procedures with isotonic approach for the selection of unknown means of normal populations and gamma populations are given. Monte Carlo comparisons on the performance of several procedures for the normal or gamma means problem were carried out in

several selected cases; these are given in Table V and Table VI at the end of Chapter I. In each case ten thousand simulations were performed. The results of this study seem to indicate that the procedures based on isotonic estimators always have superior performance, expecially, when there are more than one bad populations (in comparison with the control).

Chapter II deals with a new 'Bayes-P^*' approach about the problem of selecting a subset which contains the 'best' of k populations. Here, by best we mean the (unknown) population with the largest unknown mean. The (non-randomized) Bayes-P^* rule refers to a rule with minimum risk in the class of (non-randomized) rules which satisfy the condition that the posterior probability of selecting the best is at least equal to P^*. Given the priors of the unknown parameters, two 'Bayes-P^*' subset selection procedures ψ^B and ψ^B_{NR} (randomized and non-randomized, respectively) under certain loss functions are obtained and compared with the classical maximum-type means procedure ψ^M. The comparisons of the performance of ψ^B with ψ^B_{NR} and ψ^M, based on Monte Carlo studies, indicate that the procedure ψ^B always has higher 'efficiency' and smaller expected selected size of the selected subset. Also ψ^B appears to be robust when the true distributions are not normal but are some other symmetric distributions such as, the logistic, the double exponential, Laplace, and the gross error model (the contaminated distribution).

CHAPTER I

SELECTION PROCEDURES FOR POPULATIONS BETTER THAN A CONTROL UNDER ORDERING PRIOR

1.1. Introduction

In this chapter, three new selection procedures are given for the problem of selecting a subset which contains all populations better than a standard or control under simple or partial ordering prior. Here by simple or partial ordering prior we mean that there exist known simple or partial order relationships (defined more specifically later in Section 1.2) among unknown parameters. The procedures described do meet the usual requirement that the probabilities of a correct selection are greater than or equal to a predetermined number P*, the so-called P* condition.

Many authors have considered the problem of comparing populations with a control under different types of formulations (see Gupta and Panchapakesan (1979)). Dunnett (1955) considered the problem of separating those treatments which are better than the control from those that are worse. Gupta and Sobel (1958), Gupta (1965), Naik (1975), Broström (1977) studied the problem of selecting a subset containing all populations better than the control. Lehmann (1961) discussed similar problems with emphasis on the derivation of a restricted minimax procedure. Kim (1979), Hsiao (1979) studied the problem of

selecting populations close to a control. In all these papers it is assumed that all populations are independent and that there is no information about the order of unknown parameters. However, in many situations, we may know something about the unknown parameters. What we know is always not the prior distributions but some partial or incomplete prior information, such as the simple or partial order relationship among the unknown parameters. This type of information about the ordering prior may come from the past experiences; or it may arise in the experiments where, for example, higher dose level of some drugs always has larger effect (side-effect) on the patients.

In Section 1.2 definitions and notations used in this chapter are introduced. In Section 1.3 we consider the problem for location parameters. We propose three types of selection procedures for the cases when the control parameter is known or not known (the scale parameter may or may not be assumed known). Some equivalent forms of the procedures are given, and their properties are discussed. In Section 1.4 the problem for scale parameters of the gamma distributions is considered and three analogous selection procedures are proposed. In both Section 1.3 and 1.4 simple ordering priors are assumed and some theorems in the theory of random walks are used. In Section 1.5 a selection procedure is given for the problem of selecting all populations better than the control under partial ordering prior. Section 1.6 deals with the use of Monte Carlo techniques to make comparisons among the selection procedures proposed in Section 1.3 and those in Section 1.4, respectively.

1.2. Notations and Definitions

Suppose we have k + 1 populations $\pi_0, \pi_1, \ldots, \pi_k$. The population treatment π_0 is called the control or standard population. Assume that the random variables X_{ij} associated with $F(\cdot;\theta_i)$ and $X_{i1}, \ldots, X_{in_i}$, $i = 1,\ldots,k$, is an independent sample from π_i. Assume that we have an ordering prior of $\theta_1,\ldots,\theta_k$. First we assume that the ordering prior is the simple order, so that without loss of generality, we may assume that, $\theta_1 \le \ldots \le \theta_k$. In Section 1.5 we will consider the partial ordering prior case. Note that the values of θ_i's are unknown.

Suppose our goal is to find a non-trivial (small) subset which contains all populations with parameter larger (smaller) than the control θ_0 (known or unknown) with probability not less than a given value P^*.

The action space $\mathcal{A}$ is the class of all subsets of set $\{1, 2, \ldots, k\}$. An action A is the selection of some subset of the k populations. $i \in A$ means that π_i is included in the selected subset.

Let $\theta = (\theta_0, \theta_1, \ldots, \theta_k)$. Then the parameter space is denoted by Ω, where $\Omega = \{\theta \in \mathbb{R}^{k+1} \mid \theta_1 \le \theta_2 \le \ldots \le \theta_k;\ -\infty < \theta_0 < \infty\}$ is a subset of k + 1 dimensional Euclidean space $\mathbb{R}^{k+1}$.

The sample space is denoted by $\mathcal{X}$ where

$$\mathcal{X} = \{x \in \mathbb{R}^{n_1+\ldots+n_k},\ x = (x_{11},\ldots x_{1n_1}, x_{21},\ldots,x_{k1},\ldots x_{kn_k})\}.$$

Definition 1.2.1. A (non-randomized) selection procedure (rule) $\delta(x)$ is a mapping from $\mathcal{X}$ to $\mathcal{A}$.

A population π_i $(i = 1,\ldots,k)$ is called a good population if $\theta_i \geq \theta_0$, and we say a selection procedure δ make a correct selection (CS) if the selected subset contains all good populations. A selection procedure δ satisfies the P*-condition if

$$P_\theta(CS|\delta) \geq P^* \quad \text{for all} \quad \theta \in \Omega$$

that is

$$\inf_{\theta \in \Omega} P_\theta(CS|\delta) \geq P^*. \tag{1.2.1}$$

Let $\mathcal{A} = \{\delta | \inf_{\theta \in \Omega} P_\theta(CS|\delta) \geq P^*\}$ be a collection of all selection procedures satisfying the P*-condition.

In the sequel we will use the isotonic estimators (see Barlow, Bartholomew, Bremner and Brunk (1972)). Hence we give the following definitions and theorems.

Definition 1.2.2. Let the set $\mathcal{X}$ be a finite set. A binary relation "$\preceq$" on $\mathcal{X}$ is called a simple order if it is

(1) reflexive: $x \preceq x$ for $x \in \mathcal{X}$

(2) transitive: $x, y, z \in \mathcal{X}$ and $x \preceq y$, $y \preceq z$ imply $x \preceq z$

(3) antisymmetric: $x, y \in \mathcal{X}$ and $x \preceq y$, $y \preceq x$ imply $x = y$

(4) every two elements are comparable: $x, y \in \mathcal{X}$ imply either $x \preceq y$ or $y \preceq x$.

A partial order on $\mathcal{X}$ is a binary relation "$\preceq$" on $\mathcal{X}$, such that it is (1) reflexive, (2) transitive, and (3) antisymmetric. Thus every simple order is a partial order. We use poset $(\mathcal{X}, \preceq)$ to denote the set $\mathcal{X}$ that has a partial order binary relation "$\preceq$" on it.

Definition 1.2.3. A real-valued function f is called isotonic on poset ($\mathcal{J}$, [illegible]) if and only if (1) f is defined on $\mathcal{J}$, (2) if x, y [illegible] f(x) [illegible] f(y).

Definition 1.2.4. Let g be a real-valued function on [illegible] and let W be a given positive function on $\mathcal{J}$. A function g^* on $\mathcal{J}$ is called [illegible] regression of g with weights W if and only if:

(1) g^* is an isotonic function on poset ([illegible])

(2) $$\sum_{x \in \mathcal{J}} [g(x) - g^*(x)]^2 W(x) = \min_{f \in \mathcal{F}} \sum_{x \in \mathcal{J}} [g(x) - f(x)]^2 W(x),$$

where $\mathcal{F}$ is the class of all isotonic functions on poset [illegible].

From Barlow, et. al. (1972), (see their Theorem [illegible] and [illegible] corollary there), we have the following theorems.

Theorem 1.2.1. There exists one and only one isotonic [illegible] of g with weight W on poset [illegible].

Definition 1.2.5. A set S is convex if s_1 and [illegible] $\in S$ and [illegible] then [illegible] s_1 + (1-[illegible])$s_2 \in S$.

Definition 1.2.6. A set S is a cone if [illegible] then for any non-negative real number [illegible], [illegible] $\in S$.

Definition 1.2.7. A poset [illegible] is a lattice if [illegible] and [illegible] for any finite non-empty subset [illegible] of [illegible].

If f and g are two isotonic functions on poset [illegible], we define $f \wedge g$ and $f \vee g$ as

$$(f \wedge g)(t) = f(t) \wedge g(t) = \min(f(t), [illegible]$$

and

$$(f \vee g)(t) = f(t) \vee g(t) = \max(f(t), g(t)).$$

Then we state the following:

Theorem 1.2.2. The class [illegible] of all isotonic functions on poset [illegible] is a convex cone and a lattice.

There are some algorithms, such as the "pool-adjacent-violation algorithm (see page 13 of Barlow, et. al. (1972)) or Ayer, Brunk, [illegible], Reid and Silverman (1955) or the "up-and-down blocks" algorithm, [illegible] (1964), which show how to calculate the isotonic regression under [illegible] order.

The following max-min formulas were given by Ayer et. al. [illegible].

Theorem 1.2.3. (max-min formulas)

Assume that we have poset $(\Omega; \leq)$ where $\Omega = \{x_1, \ldots, x_k\}$ [illegible] and that function $g: \Omega \to \mathbb{R}$, then the isotonic regression g^* of g with weight W has the following formulas:

$$
\begin{aligned}
g^*(x_i) &= \max_{s \leq i} \min_{t \geq i} Av(s,t) \\
&= \max_{s \leq i} \min_{t \geq s} Av(s,t) \\
&= \min_{t \geq i} \max_{s \leq i} Av(s,t) \\
&= \min_{t \geq i} \max_{s \leq t} Av(s,t)
\end{aligned}
$$

where

$$Av(s,t) = \frac{\sum_{r=s}^{t} g(x_r) W(x_r)}{\sum_{r=s}^{t} W(x_r)}.$$

Corollary 1.2.1. $(g + c)^* = g^* + c$,

$(ag)^* = ag^*$, if a 0.

Corollary 1.2.2. $[(g^*)g + \zeta(g^*)]^* = (g^*)g^* + \zeta(g^*)$, where is a non negative function and ζ is an arbitrary function.

1.3. Proposed Selection Procedures for the Location Parameter Problem

To discuss some more general results, we assume that population $_i$ has an absolutely continuous location-scale distribution function $F(x; \mu_i, ^2) = F(\frac{x-\mu_i}{})$, where $0 < F(x) < 1$ for all x, and the values of $\mu_1, \ldots,$ and μ_k are unknown, but their ordering, say, $\mu_1 \ldots \mu_k$ is known. Note that in this case we replace in the parameter space by , all other quantities remaining the same.

Let us define the subspace $_i$ $k-i$ μ_0 $i-i+1$ for $i = 1,\ldots,k-1$ and let subspace $_k$ 0 1 and subspace $_0 =$ μ_k μ_0, then we have $= \bigcup_{i=0}^{k} {}_i$. Note that the control μ_0 could be known or unknown. If μ_0 is unknown, we assume that the distribution of population $_0$ is $F(x; \mu_0, ^2)$ and we take independent observations $X_{01},\ldots,X_{0n}$ from $_0$ and the sample space turns to $\{x \in \mathbb{R}^{n_0+\ldots+n_k} | x = (X_{01},\ldots,X_{0n_0}, X_{21},\ldots,X_{kn_k})\}$. Using the partition $\{_0,\ldots,_k\}$ of parameter space , we have

$$\inf_{\in} P(CS|) = \inf_{1 \le i \le k} \{\inf_{\in _i} P(CS|)\},$$

for any selection procedure $\in$. Hence the P*-condition is equivalent to

$$\inf_{\mu \in \Omega_i} P_\mu(CS|\delta) \geq P^*, \quad \text{for } i = 1,\ldots,k.$$

Note that $\inf_{\mu \in \Omega_0} P_\mu(CS|\delta) = 1$ for any selection procedure δ since there exist no good population in this case.

Suppose $\bar{X}_i = \bar{x}_i$ is the outcome of the sample mean of population π_i, $i = 1,\ldots,k$. Let $\mathcal{S}$ denote the set $\{\mu_1, \mu_2, \ldots, \mu_k\}$ where $\mu_1 \leq \ldots \leq \mu_k$, and let $W(\mu_i) = n_i \sigma^{-2} = w_i$, $g(\mu_i) = \bar{x}_i$, $i = 1,\ldots,k$. Then by the max-min formulas, the isotonic regression of g is g^*, where

$$g^*(\mu_i) = \max_{1 \leq s \leq i} \min_{s \leq t \leq k} \frac{\sum_{j=s}^{t} \bar{x}_j w_j}{\sum_{j=s}^{t} w_j}, \quad i = 1,\ldots,k.$$

The isotonic estimator of μ_i is denoted by $\bar{X}_{i:k}$, $i = 1,\ldots,k$ where

$$\begin{aligned} \bar{X}_{i:k} &= \max_{1 \leq s \leq i} \min_{s \leq t \leq k} \frac{\sum_{j=s}^{t} \bar{X}_j w_j}{\sum_{j=s}^{t} w_j} \\ &= \max_{1 \leq j \leq i} \{\hat{\bar{X}}_{j:k}\} \end{aligned} \tag{1.3.1}$$

where

$$\hat{\bar{X}}_{j:k} = \min\left\{\bar{X}_j, \frac{\bar{X}_j w_j + \bar{X}_{j+1} w_{j+1}}{w_j + w_{j+1}}, \ldots, \frac{\bar{X}_j w_j + \ldots + \bar{X}_k w_k}{w_j + \ldots + w_k}\right\}. \tag{1.3.2}$$

1.3.1. <u>Proposed Selection Procedure δ_1</u>

Case I. μ_0 known, common variance σ^2 known, and common sample size n.

Definition 1.3.1. We define the procedure R_1 as follows:

Step 1. Select π_i, $i = 1,\dots,k$ and stop, if

$$\bar{X}_{1:k} \ge \mu_0 - d^{(1)}_{1:k}\frac{\sigma}{\sqrt{n}},$$

otherwise reject π_1 and go to step 2.

Step 2. Select π_i, $i = 2,\dots,k$ and stop, if

$$\bar{X}_{2:k} \ge \mu_0 - d^{(1)}_{2:k}\frac{\sigma}{\sqrt{n}},$$

otherwise reject π_2 and go to step 3.

.
.
.

Step k-1. Select π_i, $i = k-1, k$ and stop, if

$$\bar{X}_{k-1:k} \ge \mu_0 - d^{(1)}_{k-1:k}\frac{\sigma}{\sqrt{n}},$$

otherwise reject π_{k-1} and go to step k.

Step K. Select π_k and stop, if

$$\bar{X}_{k:k} \ge \mu_0 - d^{(1)}_{k:k}\frac{\sigma}{\sqrt{n}},$$

otherwise reject π_k.

Here $d^{(1)}_{i:k}$'s are the smallest values such that $R_1 \in \mathcal{A}$, that is, R_1 satisfies the p*-condition.

1.3.2. On the Evaluation of $\inf_{\mu\in\Omega_i} P_\mu(CS|R_1)$ and the Value of the Constants $d^{(1)}_{1:k},\dots,d^{(1)}_{k:k}$

For any $\mu\in\Omega_i$, $1 \le i \le k$, let Z_i's i.i.d. $\Phi(\cdot;\,0,\,1)$ then

$$
\begin{aligned}
&P_\mu(CS|R_1)\\
&= P_\mu\Big(\bigcup_{j=1}^{k-i+1}\{X_{j:k} \ge \mu_0 - d^{(1)}_{j:k}\tfrac{\sigma}{\sqrt{n}}\}\Big)\\
&= P_\mu\Big(\bigcup_{j=1}^{k-i+1}\bigcup_{r=1}^{j}\{X_{r:k} \ge \mu_0 - d^{(1)}_{j:k}\tfrac{\sigma}{\sqrt{n}}\}\Big)\\
&= P_\mu\Big(\bigcup_{j=1}^{k-i+1}\bigcup_{r=1}^{j}\{Z_{r:k} + \frac{\mu_r-\mu_0}{\sigma/\sqrt{n}} \ge - d^{(1)}_{j:k}\}\Big)
\end{aligned}
$$

which is decreasing in μ_r, $r = 1,\dots,k-i+1$.

Hence

$$\inf_{\mu\in\Omega_i} P_\mu(CS|R_1) \ge P(Z_{k-i+1:k} \ge - d^{(1)}_{k-i+1:k})$$

On the other hand,

$$
\begin{aligned}
&\inf_{\mu\in\Omega_i} P_\mu(CS|R_1)\\
&\le P_{\mu^\star}\Big(\bigcup_{j=1}^{k-i+1}\{X_{j:k} \ge \mu_0 - d^{(1)}_{j:k}\tfrac{\sigma}{\sqrt{n}}\}\Big)\\
&= P(Z_{k-i+1:k} \ge - d^{(1)}_{k-i+1:k})
\end{aligned}
$$

whenever $\mu^\star = (\mu_0, -\infty,\dots,-\infty, \underbrace{\mu_0,\dots,\mu_0}_{i}) \in \Omega_i$

Thus, we have

$$\inf_{\mu\in\Omega_i} P_\mu(CS|R_1) = P(Z_{k-i+1:k} \ge - d^{(1)}_{k-i+1:k}).$$

Since $\hat{Z}_{k-i+1:k} = \min\{Z_{k-i+1},\ldots,\frac{Z_{k-i+1}+\ldots+Z_k}{i}\}$ has the same distributions as

$$\hat{Z}_{1:i} = \min\{Z_1,\ldots,\frac{Z_1+\ldots+Z_i}{i}\},$$

let

$$V_i = \hat{Z}_{1:i} \tag{1.3.3}$$

$$= \min_{1\le r\le i} \frac{1}{r}\sum_{j=1}^{r} Z_j,$$

we have

$$\inf_{\mu\in C_i} P_{\mu}(CS|s_1) = P(V_i \ge -d^{(1)}_{k-i+1:k}), \quad i = 1,\ldots,k. \tag{1.3.4}$$

Theorem 1.3.1. In case I, (μ_0 known, common known σ^2 and common sample size n), if $d^{(1)}_{k-i+1:k}$ is the solution of equation

$$P(V_i \ge -x) = P^* \tag{1.3.5}$$

where

$$V_i = \min_{1\le r\le i} \frac{1}{r}\sum_{j=1}^{r} Z_j \quad \text{and} \quad Z_i \text{ are i.i.d. } F(\cdot),$$

$i = 1,\ldots,k$ then s_1 satisfies the P^*-condition.

Proof. For any i, $1 \le i \le k$,

$$\inf_{\mu\in C_i} P_{\mu}(CS|s_1) = P(V_i \ge -d^{(1)}_{k-i+1:k}) = P^*,$$

so s_1 satisfies the P^*-condition.

Therefore, the problem of finding the $d^{(1)}_{i:k}$'s reduces to finding the distributions of $V_1,\ldots,$ and V_k. This is achieved by using some theorems in the theory of random walk.

1.3.3. Some Theorems in the Theory of Random Walk

Suppose $Y_1, Y_2, \ldots$ are independent random variables with a common distribution H not concentrated on a half-axis, i.e. $0 < P(Y_1 > 0)$, $P(Y_1 > 0) < 1$. The induced random walk is the sequence of random variables

$$S_0 = 0,\ S_n = Y_1 + \ldots + Y_n,\quad n = 1, 2, \ldots .$$

Let

$$\tau_n = P(S_1 \le 0, \ldots, S_{n-1} \le 0,\ S_n > 0) \qquad (1.3.6)$$

and

$$\tau(s) = \sum_{n=1}^{\infty} \tau_n s^n, \quad 0 \le s \le 1. \qquad (1.3.7)$$

Then we have the following theorem which was discovered by Andersen (1953). Feller (1971) gave an elegant short proof.

Theorem 1.3.2.

$$\log \frac{1}{1-\tau(s)} = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n > 0). \qquad (1.3.8)$$

Theorem 1.3.3. (Feller (1971))

Let

$$p_n = P(S_1 \le 0, \ldots, S_n \le 0),$$

then

$$p(s) = \sum_{n=1}^{\infty} p_n s^n = \frac{1}{1-\tau(s)}, \qquad (1.3.9)$$

hence

$$\log p(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n \le 0). \qquad (1.3.10)$$

By symmetry, the probabilities

$$q_n = P(S_1 \leq 0, \ldots, S_n \leq 0) \qquad (1.3.11)$$

have the generating function q given by

$$\log q(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n \leq 0). \qquad (1.3.12)$$

Note: The above two theorems remain valid if the signs [illegible] and [illegible] are replaced by $>$ and $<$, respectively.

Now, let

$$U_j' = \max_{1 \leq r \leq j} \frac{1}{r} \sum_{i=1}^{r} Z_i', \quad j = 1,2,\ldots, \qquad (1.3.13)$$

and

$$V_j' = \min_{1 \leq r \leq j} \frac{1}{r} \sum_{i=1}^{r} Z_i', \quad j = 1,2,\ldots, \qquad (1.3.14)$$

where Z_i''s are i.i.d. with absolutely continuous c.d.f. $G(\cdot)$. We would like to apply Theorem 1.3.3 to get the distribution of U_i' and V_j', $j = 1,2,\ldots$.

Remark 1.3.1. The distribution of U_j', $j = 1,\ldots,k$ for some $k \geq 1$, will be used whenever our goal is changed to select a subset containing no population with parameter smaller than the control.

Theorem 1.3.4. The generating function $q(s)$ of $P(U_j' \leq x)$, $j = 1, \ldots$ is

$$\sum_{j=1}^{\infty} s^j P(U_j' \leq x) = [illegible] \sum_{n=1}^{\infty} \frac{1}{n} s^n P(S_n [illegible] 0) [illegible]$$

where

$$S_n = \sum_{i=1}^{n} (Z_i' - x), \quad n = 1, 2, \ldots,$$

if the distribution of $Y_1 = Z_1' - x$ is not concentrated on a half-axis.

Proof. Since the distribution of random variable $Y_i = Z_i' - x$ is not concentrated on a half-axis, and Y_i's are i.i.d. let $S_r = \sum_{i=1}^{r} (Z_i' - x)$, $r = 1, \ldots, k$. Then

$$\{U_j' < x\} = \{\max_{1 \le r \le j} \frac{1}{r} S_r < 0\} = \{S_1 < 0, \ldots, S_j < 0\}.$$

By Feller's Theorem 1.3.3, we complete the proof.

Similarly, $\{V_j' > x\} = \{S_i > 0, \ i = 1, 2, \ldots, j\}$,

where

$$S_i = \sum_{r=1}^{i} (Z_r' - x).$$

Theorem 1.3.5. The generating function $p(s)$ of $P(V_j' > x)$ is

$$\sum_{j=1}^{\infty} s^j P(V_j' > x) = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} s^n P(S_n > 0)\right\}, \qquad (1.3.16)$$

if the distribution of $Y_1 = Z_1' - x$ is not concentrated on a half-axis.

Corollary 1.3.1. Both Theorem 1.3.4 and Theorem 1.3.5 hold for all x such that $0 < G(x) < 1$.

Proof. Let $Y_1 = Z_1' - x$, then

$$P(Y_1 < 0) = G(x)$$

and

$$0 < G(x) < 1,$$

hence Y_1 is not concentrated on a half-axis.

Corollary 1.3.2. Both Theorem 1.3.4 and Theorem 1.3.5 hold for all x whenever $G = \Phi$, c.d.f. of $N(0,1)$, or $G = F$ which is defined at the beginning of Section 1.3.

Proof. Followed immediately by Corollary 1.3.1.

Note that in the case of location parameter of normal population, $P(U_n' \leq -x) = P(V_n' \geq x)$.

Let

$$\Delta_j(x) = \Delta_j = P(S_j \geq 0), \quad j = 1, 2, \ldots,$$

$$a(s) = \sum_{n=1}^{\infty} \frac{s^n}{n} \Delta_n,$$

we have

$$p(s) = \sum_{j=1}^{\infty} s^j P(V_j' \geq x) = \exp(a(s)).$$

Lemma 1.3.1. $p^{(n+1)}(s) = \sum_{j=0}^{n} \binom{n}{j} p^{(j)}(s)\, a^{(n+1-j)}(s), \quad \forall n.$

Proof. Since $p'(s) = p(s) \cdot a'(s)$, the result can be proved by induction on n.

Theorem 1.3.6. Under the assumption of Theorem 1.3.5

$$P(V'_{n+1} \le x) = \frac{1}{(n+1)!} \lim_{s \to 0^+} \frac{d^{n+1} p(s)}{ds^{n+1}}$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} P(V'_j \le x) \lambda_{n-j+1}, \quad n = 0, 1, 2, \ldots \qquad (1.3.17)$$

where

$$P(V'_0 \le x) = 1, \ \forall x.$$

Proof. By Lemma 1.3.1, we have

$$P(V'_{n+1} \le x) = \frac{1}{(n+1)!} \lim_{s \to 0^+} p^{(n+1)}(s)$$

$$= \sum_{j=0}^{n} \frac{1}{(n+1)!} \frac{n!}{j!(n-j)!} p^{(j)}(0) \, [(n-j)! \, \lambda_{n+1-j}]$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} \frac{p^{(j)}(0)}{j!} \lambda_{n+1-j}$$

$$= \frac{1}{n+1} \sum_{j=0}^{n} P(V'_j \le x) \lambda_{n+1-j}.$$

Similarly, we have

$$P(U'_{n+1} \le x) = \frac{1}{n+1} \sum_{i=0}^{n} P(U'_{n-i+1} \le x) \, P(S_i \le 0). \qquad (1.3.18)$$

1.3.4. Limiting Distributions of U'_n and V'_n

Let $F_n(x) = P(U'_n \le x)$ and $F_\infty(x)$ denote the limiting distribution function as $n \to \infty$ of U'_n. Suppose the distribution of random variable $Y_1 = Z'_1 - x$ is not concentrated on a half axis, then we have

$$1 - F_n(x) = P(S_1 > 0) + \sum_{r=2}^{n} P(S_1 \le 0, \ldots, S_{r-1} \le 0, S_r > 0),$$

$$1 - F_\infty(x) = \lim_{s \to 1^-} \phi(s),$$

and apply Andersen-Feller Theorem 1.3.2, we have

$$F_\infty(x) = \exp\{- \sum_{r=1}^{\infty} \frac{1}{r} P(S_r > 0)\}. \qquad (1.3.19)$$

Similarly,

$$G_\infty(x) = \exp\{- \sum_{r=1}^{\infty} \frac{1}{r} P(S_r \ge 0)\} \qquad (1.3.20)$$

where

$$G_\infty(x) = P(V'_\infty < x).$$

Let

$$G_\infty(- d^{(1)}_{1:\infty}) = P^*. \qquad (1.3.21)$$

If Z_i, $i = 1,\ldots,k$, are independent identically distributed $N(0,1)$, then we can use the recurrence formula of Theorem 1.3.6 to solve the equations $P(V_i > - d_{k-i+1:k}) = P^*$, $i = 1,\ldots,k$. Hence in Case I, $F_i(x) = \Phi(- x\sqrt{j})$.

Remark 1.3.2. From formula (1.3.4) we know that $d^{(1)}_{k-i+1:k}$ $(i = 1,\ldots,k)$ does not depend on k. And we have $d^{(1)}_{k-i+1:k} = d^{(1)}_{1:i}$. These values for $k = 1\ (1)\ 6,\ 10,\ \infty$ and $P^* = .99,\ .975,\ .95,\ .925,\ .9,\ .85,\ .8,\ .75,\ .7,\ .65$ are tabulated in Table I.

1.3.5. Some Other Forms of Selection Procedure 1

Lemma 1.3.2. $d^{(1)}_{1:i}$ is increasing in i.

Proof. By Remark 1.3.2 and the fact

$$V_{i+1} = \min\left(V_i, \frac{iV_i + Z_{i+1}}{i+1}\right).$$

Lemma 1.3.3. If c_j, $1 \le j \le i \le k$ is decreasing in j, then

$$\bigcup_{j=1}^{i} \{X_{j:k} \ge -c_j\} = \bigcup_{j=1}^{i} \{X_{j:k} \ge -c_j\}$$

Proof.

$$\bigcup_{j=1}^{i} \{X_{j:k} \ge -c_j\} \subset \bigcup_{j=1}^{i} \{X_{j:k} \ge -c_j\},$$

since

$$X_{j:k} \le \hat{X}_{j:k}, \quad 1 \le j \le k.$$

On the other hand, if

$$X_{r:k} \ge -c_r \quad \text{for some } r, \quad 1 \le r \le i$$

then

$$X_{s:k} \ge -c_r \quad \text{for some } s, \quad 1 \le s \le r,$$

since

$$X_{r:k} = \max_{1 \le s \le r} \{X_{s:k}\}.$$

Because c_j is decreasing in j, this implies $X_{s:k} \ge -c_s$ for some s, $1 \le s \le r$.

Hence we have

$$\bigcup_{j=1}^{i} \{X_{j:k} > \ldots - c_j\} \subset \bigcup_{j=1}^{i} \{X_{j:k} > \ldots - c_j\},$$

therefore the lemma is proved.

Definition 1.3.2. We define a selection procedure R_1' by replacing the inequality in the ith step of procedure R_1 by the inequality

$$X_{i:k} \geq \ldots - d'_{i:k} \frac{\ldots}{\sqrt{n}}, \quad i = 1,\ldots,k$$

where $d'_{1:k},\ldots,d'_{k:k}$ are the smallest values such that R_1' satisfies the P^*-condition.

Theorem 1.3.7. The selection procedure R_1 and R_1' are identical iff $d^{(1)}_{i:k} = d'_{i:k}$, $i = 1,\ldots,k$.

Proof. For any i, $1 \leq i \leq k$, by Theorem 1.3.1

$$P^* = \inf_{\in \Omega_i} P(CS|R_1) = P(Z_{k-i+1:k} \ldots - d^{(1)}_{k-i+1:k}).$$

On the other hand, using the same arguments as Section 1.3.2, we have

$$P^* = \inf_{\in \Omega_i} P(CS|R_1') = P(Z_{k-i+1:k} \ldots - d'_{k-i+1:k}).$$

Hence we have $d^{(1)}_{i:k} = d'_{i:k}$, $i = 1,\ldots,k$.

Since $X_{1:k} \ldots X_{i:k}$, the first step of R_1 and R_1' are identical. For $i = 2,\ldots,k$, the event

$$\{\text{select } \pi_i, \ldots, \pi_k\} = \bigcup_{j=1}^{i} (X_{j:k} \geq \theta_0 - d_{i:k}^{(1)}\ldots)$$

$$\bigcup_{j=1}^{i} (X_{j:k} \geq \theta_0 - d_{i:k}^{(1)}\ldots_{,n})$$

$$= \{\text{select } \pi_i, \ldots, \pi_k \mid \delta_1\}$$

by Lemma 1.3.2 and Lemma 1.3.3. Hence selection procedures δ_1 and δ_1' are identical.

1.3.6. Some Other Proposed Selection Procedures (δ_2, δ_3)

In Case I, we proposed some other selection procedures:

Definition 1.3.3. We define a selection procedure δ_2 by

δ_2: Select π_i if and only if $X_{i:k} \geq \theta_0 - d_{i,n}$, $i = 1, \ldots, k$

where d is the smallest value such that δ_2 satisfies the P*-condition.

Theorem 1.3.8. Under assumptions of Case I, and selection procedure δ_2, if we select population π_i, then we will select populations π_j, for all $j > i$.

Proof. Since $X_{i:k} \leq X_{j:k}$ for all $j > i$.

Evaluation the Value d of δ_2

For any i, $1 \leq i \leq k$, we have

$$\inf_{\Omega_i} P_{\theta}(CS \mid \delta_2) = \inf_{\Omega_i} P_{\theta}(X_{k-i+1:k} \geq \theta_0 - d_{i,n})$$

$$= P(Y_i \geq -d)$$

by the same argument for selection procedure R_1 and here

$$V_i = Z_{1:i} = \min_{1 \le r \le i} \frac{1}{r} \sum_{j=1}^{r} Z_j .$$

We need the constant d such that $P(V_i \ge -d) \ge P^*$ holds for all i, $1 \le i \le k$. By Lemma 1.3.2 we have $d = d^{(1)}_{1:k}$. Hence we have the following theorem.

Theorem 1.3.9. Selection procedure R_2 satisfies the P^*-condition with $d = d^{(1)}_{1:k}$.

Corollary 1.3.3. If S_1 and S_2 are the selected subsets associated with selection procedures R_1 and R_2, respectively, then $S_1 \subseteq S_2$.

Proof. Proof follows from Lemma 1.3.2.

Definition 1.3.4. The procedure R_3 is defined as follows:

Step 1. Select π_i, $i \ge 1$ and stop, if

$$\bar{X}_1 \ge \mu_0 - d_1 \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_1 and go to step 2.

Step 2. Select π_i, $i \ge 2$ and stop, if

$$\bar{X}_2 \ge \mu_0 - d_2 \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_2 and go to step 3.

.
.
.

Step k-1. Select π_i, $i \ge k - 1$ and stop, if

$$\bar{X}_{k-1} \ge \mu_0 - d_{k-1} \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_{k-1} and go to step k.

Step k. Select π_k and stop, if

$$\bar{X}_k \ge \mu_0 - d_k \frac{\sigma}{\sqrt{n}},$$

otherwise reject π_k.

Here $\bar{X}_j = \max\{X_1, \ldots X_j\}$ and d_i's are the smallest values such that R_i satisfies the P*-condition.

Evaluation of d_i's

For any i, $1 \le i \le k$,

$$\inf_{\mu \in \Omega_i} P_\mu(CS|R_3) = \inf_{\mu \in \Omega_i} P_\mu\left(\bigcup_{j=1}^{k-i+1} \{X_j \ge \mu_0 - d_j \frac{\sigma}{\sqrt{n}}\}\right)$$

$$= P_\mu(Z_{k-i+1} \ge - d_{k-i+1} \frac{\sigma}{\sqrt{n}})$$

$$= P(Z_{k-i+1} \ge - d_{k-i+1})$$

$$= P^*, \quad Z_i \sim F(\cdot;\ 0,\ 1).$$

This implies $d_{k-i+1} = d$ for all i, and

$$d = - F^{-1}(1-P^*),$$

$$= F^{-1}(P^*), \text{ if F is symmetric}$$

$$= \Phi^{-1}(P^*), \text{ if } X_i \sim N(\mu_i, \sigma^2/n).$$

Similar to the selection procedure R_1, we have the following theorem:

Theorem 1.3.10. Selection procedure R_3 satisfies the P*-condition with $d_i = - F^{-1}(1-P^*)$.

Definition 1.3.5. Selection procedures δ_2 is defined as follows:

Step 1. Select π_i, $i = 1$ and stop, if

$$X_1 \ [illegible] \ X_0 - d_{[illegible],n},$$

otherwise reject π_1 and go to step 2.

Step 2. Select π_i, $i = 2$ and stop, if

$$X_2 \ [illegible] \ X_0 - d_{[illegible],n},$$

otherwise reject π_2 and go to step 3.

.
.
.

Step k-1. Select π_i, $i = k - 1$ and stop, if

$$X_{k-1} \ [illegible] \ X_0 - d_{[illegible],n},$$

otherwise reject π_k.

Here

$$d \ [illegible] = -F^{-1}(1-P^*)$$

$$= F^{-1}(P^*) \text{ if } F \text{ is symmetric.}$$

Theorem 1.3.11. The selection procedures δ_2 satisfies the P^* condition.

Proof. For any i, $1 \le i \le k$,

$$\inf_{\Omega_i} P_{\theta}(CS \mid \delta_2) = P([illegible]_{k-1} [illegible] - D) = P^*.$$

Theorem 1.3.12. The selection procedure δ_1 and δ_2 are identical.

Proof. The proof is simple hence it is omitted.

The following procedure δ_4 was given by Gupta and Sobel (1958),
without assuming any ordering prior:

Definition 1.3.6. The selection procedure δ_4 is defined as follows:

δ_4: Select π_i if and only if $X_i \geq X_0 - d_{\cdot n}$, $i = 1,\ldots,k$

where d is the smallest constant such that δ_4 satisfies the P*-condition.

It was shown that the value d is determined by the equation

$$F(-d) = 1 - P^{*\frac{1}{k}}$$

or

$$F(d) = P^{*\frac{1}{k}} \text{ if } F \text{ is symmetric.}$$

1.3.7. A Dual Problem

We start with the same assumptions as in Section 1.3.1 Case 1, but change our goal to select a subset which contains no bad populations; the definition of a correct selection (CS) will now be changed to select a subset that contains no bad populations and the P*-condition will be defined based on this new definition of correct selection (CS).

In location parameter case, this problem is a dual problem of the original problem, namely, "select a subset which contains all good populations under ordering prior assumption".

One method to solve this problem is that, first, change the signs of all statistics and the control to opposite sign; then use a procedure for selecting a subset which contains all "new good" populations.

where the "new good" populations are the "old bad" populations before changing signs; finally, reject the selected subset and keep the remainders as the desired selected subset. Let [illegible]$_i$, i = 1, 2, 3, 4 denote the above procedure which corresponds to [illegible]$_i$, i = 1, 2, 3, 4, respectively.

Theorem 1.3.13. The selection procedure [illegible]$_i$, i = 1, 2, 3, 4 satisfies the P*-condition in which the correct selection (CS) means that it selects a subset which contains no bad population.

Proof. Given P* and observations, for any selection procedure [illegible]$_i$, i = 1, 2, 3, 4, after changing the signs of all associated statistics, the probability that the selected subset S contains all "new good" populations is not less than P*. If we reject the selected subset S, then the complement subset S^c of S contains any "new good" populations with probability less than 1-P*, but the "new good" populations are the originally bad populations so what we have is that the subset S^c contains any originally bad population with probability less than 1-P*, in other words, subset S^c contains no bad population with probability greater than or equal to P*. Since this is true for all arbitrary true configurations, we have completed the proof.

Remark 1.3.3. It is easy to see that the value $d_{i:k}$ [illegible] of [illegible] which was used by [illegible] in the ith step is determined by the equation

$$P(U_{k-i+1} \leq d_{i:k}(\ldots)) = P^*,$$

where

[illegible]

If the distribution F is symmetric, then

$$d_{i:k}(\cdot_1) = d^{(1)}_{i:k}.$$

1.3.8. Some Proposed Selection Procedures $R^{(2)}_i$, i = 1, 2, 3, 4 When $_0$ is Unknown

Case II. $_0$ unknown, common ^{2} known, common sample size n.

Definition 1.3.7. We define a selection procedure $R^{(2)}_1$ by replacing the inequalities

$$X_{i:k} \quad _0 - d^{(1)}_{i:k} \quad /\sqrt{n}, \quad i = 1,\dots,k$$

in procedure R_1 (Definition 1.3.1) with

$$X_{i:k} \quad X_0 - d^{(2)}_{i:k} \quad /\sqrt{n}, \quad i = 1,\dots,k, \text{ respectively.}$$

Here $X_0 = \sum_{i=1}^{n} X_{0i}/n$, $d^{(2)}_{i:k}$, i = 1,...,k are the smallest constants such that the selection procedure $R^{(2)}$ satisfies the P*-condition.

Similar to the Case I, we have the following theorem:

Theorem 1.3.14. For any i, 1 i k, $d^{(2)}_{k-i+1:k}$ is determined by the equation

$$\int F(V_1 \quad t - d^{(2)}_{k-i+1:k}) \, dH(t) \quad P^*, \qquad 1, \dots$$

It is easy to see that $d^{(2)}_{k-i+1:k} = d^{(2)}_{1:i}$. The following theorem gives us an identical form of the selection procedure $\delta_1^{(2)}$.

Theorem 1.3.15. The selection procedure $\delta_1^{(2)}$ will not be changed if the statistics $X_{i:k}$, $i = 1,\ldots,k$, are replaced by $X'_{i:k}$, $i = 1,\ldots,k$, respectively.

Proof. The proof is the same as that in Case I and hence it is omitted.

The values $d^{(2)}_{1:i}$, $i = 1,\ldots,k$ are tabulated in Table II for $k = 1\ (1)\ 6,\ 8,\ 10,$ and $P^* = .99, .975, .95, .925, .90, .85, .80, .75, .70, .65$.

Similar to the Case I, we propose a selection procedure $\delta_2^{(2)}$ as follows:

Definition 1.3.8. We define a selection procedure $\delta_2^{(2)}$ by

$\delta_2^{(2)}$: Select π_i if and only if $X_{i:k} \geq X_0 - d/\sqrt{n}$, $i = 1,\ldots,k$

where d is the smallest value such that $\delta_2^{(2)}$ satisfies the P^* condition. Then, similar to Theorem 1.2.9 we have:

Theorem 1.3.16. Under assumptions of Case II, the selection procedure $\delta_2^{(2)}$ satisfies the P^*-condition with $d = d^{(2)}_{1:k}$.

Next, we define a selection procedure $\delta_3^{(2)}$ which is similar to $\delta_2^{(2)}$ but replace μ_0 by $\bar{X}_0$, the sample mean of population π_0.

Definition 1.3.9. The selection procedure $R_3^{(2)}$ is defined by replacing $X_i \geq X_0 - d_i \sqrt{\;}/\sqrt{n}$ in R_3 (Definition 1.3.4) by $X_i \geq X_0 - d_i' \sqrt{\;}/\sqrt{f}$, $i = 1,\ldots,k$ where $d_1',\ldots,d_k'$ are the smallest values such that $R_3^{(2)}$ satisfies the P*-condition.

Similar to Theorem 1.3.10 we have:

Theorem 1.3.12. The selection procedure $R_3^{(2)}$ satisfies the P*-condition with $d_i = d$, $i = 1,\ldots,k$ where d is determined by the equation

$$\int_{-\infty}^{\infty} [1 - F(t-d)]dF(t) = P^*, \qquad \text{[illegible]}$$

$$\int_{-\infty}^{\infty} F(d-t)dF(t) = P^*, \text{ if } F \text{ is symmetric.}$$

And $R_3^{(2)}$ will not be changed if the statistics X_i is replaced by $\bar{X}_i$, the sample mean of population π_i for $i = 1,\ldots,k$.

The following selection procedure $R_4^{(2)}$ was proposed by Gupta and Sobel (1958):

Definition 1.3.10. The selection procedure $R_4^{(2)}$ is defined by

$R_4^{(2)}$: Select π_i if and only if $X_i \geq X_0 - d \sqrt{\;}/\sqrt{n_i}$, $i = 1,\ldots,k$

where d is determined by the following equation if F is normal distribution:

$$\int_{-\infty}^{\infty} \prod_{i=1}^{k} [F(t \sqrt{n_i/n_0} + d)]f(u)du = P^*. \qquad \text{[illegible]}$$

For the special case $n_i = n$ ($i = 0, 1, \ldots, k$) [illegible]

$$\int_{-\infty}^{\infty} F^k(t+d) f(t)\,dt \; [illegible] \; P^*. \qquad (1. [illegible])$$

If F is normal distribution N(0,1), the tables of d values [illegible] isfying the Equation (1.3.25) for several values of P* are given in Bechhofer (1954) for k [illegible] and in Gupta (19[illegible]) for [illegible]

1.3.9. Some Proposed Selection Procedures [illegible], i = 1, [illegible]

When Common Variance [illegible] is Unknown

Case III. [illegible] known, common variance [illegible] unknown, n_i [illegible]

In this case, we assume that [illegible] which is the [illegible] N(0,1).

Definition 1.3.11. We define the selection procedure [illegible] by replacing the inequalities

$$[illegible]_{i:k} \; [illegible] \; - d^{(1)}_{i:k} \; [illegible], \quad i = 1, \ldots, k$$

in procedure [illegible] (Definition 1.3.1) by

$$[illegible]_{i:k} \; [illegible] \; - d^{(3)}_{i:k} \; [illegible], \quad i = 1, \ldots, k, \text{ respectively,}$$

where $d^{(3)}$'s are the smallest values such that [illegible] satisfies the P*-condition, S^2 denotes the [illegible] estimator of [illegible] based on [illegible] k(n-1), that is

$$[illegible]$$

Note that [illegible] has the chi-square distribution [illegible] with [illegible] degrees of freedom.

By using similar arguments as in Case I, we have:

Theorem 1.3.18. The equation which determines the constant $d_{k-i+1:k}^{(s)}$ is

$$P(V_i \; [illegible] \; d_{k-i+1:k} \, S) = P^* \qquad [illegible]$$

or

$$\int_0^{\infty} P(V_i \; [illegible] \; d_{k-i+1:k} \, y) q(y) dy = P^* \qquad [illegible]$$

where

$$V_i = \min_{1 \le r \le i} \frac{1}{r} \sum_{j=1}^{r} Z_j$$

and $q(y)$ is the density of [illegible].

We can rewrite Formula (1.3.28) as

$$\int_0^{\infty} P(V_i \; [illegible] \; d_{k-i+1:k}^{(s)} \sqrt{[illegible]}) \, [illegible] = P^*$$

or

$$\int_0^{\infty} P(V_i \; [illegible] \; d_{k-i+1:k}^{(s)} \, [illegible]) \, \frac{[illegible]^{\,[illegible]-1} e^{-t}}{\Gamma(\frac{[illegible]}{2})} \, dt = P^*. \qquad [illegible]$$

Remark 1.3.4. The values of $d^{(3)}_{k-i+1:k}$, $i = 1,\ldots,k$ depend on [illegible], hence $d^{(3)}_{k-i+1:k} \neq d^{(3)}_{1:i}$.

By using Rabinowitz and Weiss table (1959) (with $n =$ [illegible]), we have evaluated and tabulated the values of $d^{(3)}_{k-i+1:k}$, $i = 1,\ldots,k$, in Table III, for $k = 2\ (1)\ 6$, $P^* =$ .99, .975, .95, [illegible], .80, and .75 with common sample size $n =$ [illegible], 5, [illegible], and 21.

For $k =$ 6 and $n =$ 21, we can use $d^{(3)}_{1:i}$ as an approximation of $d^{(3)}_{k-i+1:k}$.

Definition 1.3.1[illegible]. We define the selection procedure [illegible] by

[illegible]: Select [illegible] if and only if $X_{i:k}$ [illegible]

where S is defined as in procedure [illegible], and $d^{(3)}$ is the smallest constant such that [illegible] satisfies the P*-condition.

As before, it can be shown that [illegible] $d^{(3)}_{1:k}$.

Remark 1.3.5. Theorem 1.3.[illegible] still holds for Case III, i.e. the selection procedure [illegible] will not be changed if we replace the [illegible] statistics [illegible] by [illegible], respectively. But this is not [illegible] true for selection procedure [illegible].

Definition 1.3.1[illegible]. The selection procedure [illegible] is defined to have the same form as procedure [illegible] except that the inequality [illegible] ith step of procedure [illegible] is replaced by

[illegible]

The proof of the following theorem uses the same arguments as that in Case I, hence it is omitted.

Theorem 1.3.19. The equation which determines the constant d of selection procedure $\delta_3^{(3)}$ is

$$\int_0^\infty \Phi(yd)q_\nu(y)dy = P^*. \qquad (1.3.31)$$

Gupta and Sobel (1958) gave a selection procedure $\delta_4^{(3)}$ in this case. It is as follows:

$\delta_4^{(3)}$: Select π_i if and only if $\bar{X}_i \geq \mu_0 - d \frac{S}{\sqrt{n_i}}$, $i = 1,\ldots,k$

and the equation which determines d is

$$\int_0^\infty \Phi^k(yd)q_\nu(y)dy = P^*. \qquad (1.3.32)$$

1.3.10. Some Proposed Selection Procedures $\delta_i^{(4)}$, $i = 1, 2, 3, 4$ When Both Control μ_0 and Common Variance σ^2 are Unknown

Case IV. μ_0 unknown, common variance σ^2 unknown and common sample size n.

We assume that in this case distribution F is the c.d.f. $N(0,1)$, and denoted by Φ. We replace μ_0 in each selection procedure $\delta_j^{(2)}$ by $\bar{X}_0$, $1 \leq j \leq 4$, and get four procedures $\delta_j^{(4)}$, $1 \leq j \leq 4$, respectively. Let $\chi^2_\nu(t)$ denote the c.d.f. of the chi-square distribution with $k(n-1)$ degrees of freedom. The constant $d_{k-i+1:k}^{(4)}$, $i = 1,\ldots,k$, of procedure $\delta_1^{(4)}$ is determined by

$$\int_0 \int_{-} P(Y_i \quad u - d^{(4)}_{k-i+1:k} \ldots) \ldots$$

The constant d of procedure $R_2^{(4)}$ is

$$d = d^{(4)}_{1:k}.$$

The constants d of procedures $R_3^{(4)}$ and $R_4^{(4)}$ are determined by

$$\int_0 \int_{-} \Phi^r(u - td) \ldots (u) d \ldots^2(t) \quad P^*$$

with $r = 1$ and k, respectively, and their values for selected values of P^*, k and are given in Gupta and Sobel (1957) and ...

1.3.11. Properties of the Selection Procedures

Under simple ordering prior, it is natural to require that ... selection procedure is order-preserving as defined below:

Definition 1.3.14. A selection procedure is order-preserving if it selects π_i with parameter ..., and if ..., then it also selects ... Procedure is weak order-preserving or monotone if

P($_i$ is selected) P($_j$ is selected) whenever $_i$ ...

It is easy to see that any order-preserving selection procedure is weak order-preserving, but the converse is not true.

Now, let $R_i^{(1)} = R_i$, $i = 1, 2, 3, 4$.

Theorem 1.3.20. The selection procedures $R_1^{(i)}$, ... and ... are order-preserving and procedure $R_4^{(i)}$ is monotone, for $i = 1, \ldots$

Proof. The proof follows immediately from the definitions of the pro-
cedures.

Given observations $X = x = (x_0, \ldots, x_k)$ where x_i is the sample mean of population [illegible]$_i$, $i = 1, \ldots, k$, and x_0 = [illegible]$_0$ if [illegible]$_0$ is known, otherwise x_0 is the sample mean of population [illegible]$_0$. Let

$$[\text{illegible}]_i(x, [\text{illegible}]) = P([\text{illegible}]_i \text{ included in the selected subset } [\text{illegible}])$$

for $i = 1, \ldots,$ [illegible]

Definition 1.3.16. A selection procedure [illegible] is called translation-invariant if for any $x \in \mathbb{R}^{k+1}$, $c \in \mathbb{R}$

$$[\text{illegible}]_i(x_0 + c, x_1 + c, \ldots, x_k + c; [\text{illegible}]) = [\text{illegible}]_i(x_0, \ldots, x_k; [\text{illegible}]) \quad i = 1, \ldots, [\text{illegible}]$$

Theorem 1.3.21. The selection procedures $[\text{illegible}]_1^{(i)}$, $[\text{illegible}]_2^{(i)}$, $[\text{illegible}]_3^{(i)}$ and $[\text{illegible}]_4^{(i)}$ are translation-invariant for $i = 1, 2, 3, 4$.

Proof. By Corollary 1.3.1 the isotonic regression is a linear operator. On the other hand,

$$\frac{\sum_{j=1}^{n} (x_{ij} + c)}{n} = \frac{\sum_{j=1}^{n} x_{ij}}{n} + c,$$

hence we have the result.

Expected Number (Size) of Bad Populations in the Selected Subset

Suppose the control [illegible]$_0$ is known or we have common sample size [illegible] and common known variance [illegible]; without loss of generality, we assume that [illegible]$_0$ = 0 and [illegible] = 1. Let [illegible] denote the expected number of bad populations in the selected subset in using the selection [illegible]

, then for any j, $0 \le j \le k$,

$$\sup_{\mu \in \Omega_{k-j}} E(S' \mid R_1)$$

$$= \sup_{\mu \in \Omega_{k-j}} \sum_{r=1}^{j} P\left(\bigcup_{\ell=1}^{r} Z_{\ell:k} - d_{1:k}^{(1)} \right)$$

$$= \sum_{r=1}^{j} P\left(\bigcup_{\ell=1}^{r} Z_{\ell:j} - d_{1:k}^{(1)} \right) \qquad (1.3.35)$$

On the other hand, for procedure R_2

$$\sup_{\mu \in \Omega_{k-j}} E(S' \mid R_2) = \sum_{r=1}^{j} P\left(\bigcup_{\ell=1}^{r} Z_{\ell:j} - d_{1:k}^{(1)} \right) \qquad (1.3.36)$$

Formula (1.3.36) is increasing in j and is greater than or equal to Formula (1.3.35), since

$$d_{j:k}^{(1)} = d_{1:k-j+1}^{(1)} \le d_{1:k}^{(1)}.$$

Therefore, we have the following theorem.

Theorem 1.3.22. For any i, $0 \le i \le k$

$$\sup_{\mu \in \Omega_i} E(S'' \mid R_2) \ge \sup_{\mu \in \Omega_i} E(S'' \mid R_1),$$

$$\sup_{\mu \in \Omega} E(S'' \mid R_2) = \sup_{\mu \in \Omega_0} E(S'' \mid R_2).$$

Theorem 1.3.23. In Section 1.3.1, Case 1, for any [illegible]

$$\sup_{\mu \in \Omega_{k-j}} E(S' \mid R_2) \; [illegible] \; q(1-q^{j})/P^* \qquad [illegible]$$

where $q = 1 - P^*$.

Proof.

$$\sup_{\Omega_{k-j}} E(S'_3)$$

$$= \sup_{\Omega_{k-j}} \sum_{i=1}^{j} P(\text{select } \pi_i | R_3)$$

$$= \sup_{\Omega_{k-j}} \sum_{i=1}^{j} P(\max_{1 \le r \le i} \delta_r - d)$$

$$= \sum_{i=1}^{j} (1 - \prod_{r=1}^{i} F(-d))$$

$$= j - \sum_{i=1}^{j} q^i$$

$$= j - q(1-q^j)/P^*$$

where $q = (1-P^*)$.

Theorem 1.3.24. $\sup_{\Omega_{k-j}} E(S'_3)$ is increasing in j, hence

$$\sup_{\Omega_{k-j}} E(S'_3) = \sup_{\Omega_0} E(S'_3) = k - q(1-q^k)/P^*. \quad (1.3.\ldots)$$

Proof. Since the function

$$f(x) = x - ab^{x+1}$$

is increasing in x, for $0 < a < 1$, $0 < b < 1$, and $0 \le x$.

In Case I of Gupta (1965) showed that

$$\sup_{\Omega} E(S'_4) = kP^{*\frac{1}{k}}. \quad 1.\ldots$$

When the ordering prior among the unknown parameters [illegible] known, we can use the selection procedure of Gupta and [illegible] (19[illegible]) [illegible] the ordering of the sample means as the ordering of unknown [illegible] and apply the selection procedure which [illegible] [illegible] ing prior. With the latter approach, the substitution [illegible] isotonic regression of the sample means [illegible] to the [illegible] sample means, and that the selection procedures [illegible] $j = 2, 3$) are the same as $R_{4}^{(i)}$ ($i = 1, 2, [illegible], 4$), [illegible] selection procedures $R_{1}^{(i)}$, $i = 1, 2, 3, 4$ [illegible] respectively, which are equivalent to the procedures [illegible] (1975) and Broström (1977), independently.

It has been proved in some quite general situations [illegible] by using Monte Carlo technique in some selected cases [illegible] and Broström (1977), separately, that [illegible] is slightly better than R_4. The values $d_{i:k}^{(j)}$ in the ith step of the procedure [illegible], $j = 1, 2, 3, 4$, are given by Broström (1977) as follows:

$$d_{i:k}^{(1)} = [illegible]^{-1}[1 - (P^*)^{\frac{1}{i}}], \qquad [illegible]$$

$$\int_{-\infty}^{\infty} [illegible]^{i}([illegible] + d_{i:k}^{([illegible])})[illegible] = P^*, \qquad [illegible]$$

$$\int_{0}^{\infty} [illegible]^{i}([illegible] d_{i:k}^{([illegible])})[illegible] = P^*, \qquad [illegible]$$

and

$$\int_{0}^{\infty}\int_{-\infty}^{\infty} [illegible]^{i}([illegible] + y d_{i:k}^{(4)})[illegible] q([illegible])dx\,dy = P^* \qquad [illegible]$$

where q [illegible] is the density [illegible].

1.1. Selection Procedures for Scale Parameters of Gamma Populations

Suppose we have $k+1$ independent populations [illegible]$,\ldots,$[illegible]. The [illegible] ulation π_i has a gamma density function

$$g(\cdot;\ \theta_i,\ \alpha_i) = \frac{1}{\Gamma(\alpha_i)\ \theta_i^{\alpha_i}}\ x^{\alpha_i - 1} e^{-x/\theta_i},\ \ldots \qquad (1.1)$$

$i = 1,\ldots,k$. The ordering prior of [illegible]$,\ldots,$ and θ_k [illegible], say, $\theta_1 \le \theta_2 \le \ldots \le \theta_k$. Note that the values of [illegible]$,\ldots,$ θ_k are unknown, α_i's are known.

In this section we define population π_i, $i = 1,\ldots,k$, [illegible] population if the scale parameter θ_i [illegible]. Let [illegible] $1,\ldots,$[illegible], then the parameter space is denoted by [illegible], where

$$\Omega \subset \mathbb{R}^{k+1}\ \theta\ [illegible]_1\ \ldots\ [illegible]_k\ [illegible]$$

is a subspace of $(k+1)$-dimension Euclidean space $\mathbb{R}^{k+1}$.

Suppose we have independent observations X_{ij} $(j = 1,\ldots,n_i)$ from population π_i, $(i = 1,\ldots,k)$. Let [illegible]$_i = n_i$[illegible]$_i$, then

$$X_i = \sum_{j=1}^{n_i} X_{ij}/n_i \text{ has density } g(\cdot;\ \theta_i,\ [illegible]_i)$$

and

$$X_i/[illegible]_i \text{ has density } g(\cdot;\ \theta_i,\ 1/n_i\ [illegible])$$

Suppose our goal is to select a subset which contains all [illegible] populations under the ordering prior with probability greater than or equal to P^*, a predetermined value between zero and one.

Let [illegible]$_i$, $i = [illegible],\ldots,k$, be the subspace of parameter space [illegible] such that [illegible] $\bigcap_{i=0}^{k}$ [illegible]$_i$ where

$$_i = \ldots (\ldots_i \ldots \mu_0 \ldots _{i+1}), \quad \text{if } i = 1, \ldots, \text{ or } k-1,$$

$$\ldots (\ldots_k \ldots _0), \quad \text{if } i = k,$$

$$= \{\ldots (\ldots_0 \ldots _1\}, \quad \text{if } i = 0.$$

1.4.1. Proposed Selection Procedures π_i, $i = 6, 7, 8,$ [illegible]

Case I. Control μ_0 known and common sample size n.

Definition 1.4.1. The selection procedure R_6 is defined as follow

Step 1. Select π_i, $i \le k$ and stop, if

$$X_{k:k} \ge c_{k:k}\,\mu_0,$$

otherwise reject π_k and go to step 2.

Step 2. Select π_i, $i \le k-1$ and stop, if

$$X_{k-1:k} \ge c_{k-1:k}\,\mu_0,$$

otherwise reject π_{k-1} and go to step 3.

.
.
.

Step k-1. Select π_i, $i \le 2$ and stop, if

$$X_{2:k} \ge c_{2:k}\,\mu_0,$$

otherwise reject π_2 and go to step k.

Step k. Select π_1 and stop, if

$$X_{1:k} \ge c_{1:k}\,\mu_0,$$

otherwise reject π_1.

Here $c_{i:k}$ ([illegible] 1), i [illegible] 1,...,k are the smallest values [illegible] that the procedure R_6 satisfies the P*-condition.

Theorem 1.4.1. Assume we have common sample size n and [illegible], i [illegible] 1,...,k, and the constant $c_{i:k}$ ([illegible] 1) is determined [illegible]

$$P(U_i \; [illegible] \; c_{i:k}) \; [illegible] \; P^*, \quad i = 1,\ldots,k \qquad (1.4.\ldots)$$

where

$$U_i = \max_{1 \le s \le i} \left(\frac{Y_s + \ldots + Y_i}{i-s+1} \right) \qquad (1.4.\ldots)$$

and Y_i are i.i.d. with density

$$g(\cdot; \; [illegible], \tfrac{1}{n}), \text{ and c.d.f. } G(\cdot; \; [illegible], \tfrac{1}{n})$$

then the procedure R_6 satisfies the P*-condition.

Proof. For any i, 1 [illegible] i [illegible] k, if the unknown true [illegible] that is there are i good populations, then, under the procedure [illegible],

$$\inf_{\Omega_i} P(CS \mid R_6)$$

$$\inf_{\Omega_i} P\left(\bigcap_{j=1}^{k} [illegible] \; c_{i:k} \; [illegible] \right)$$

$$\inf_{\Omega_i} P\left(\bigcap [illegible] \right)$$

$$\inf_{\Omega_i} [illegible]$$

$$= \inf_{\theta \in \Theta_i} P_\theta\left(\bigcup_{j=i}^{k} \left(\max_{1 \le s \le j} \min_{j \le t \le k} \frac{Y_s - \theta_s + \ldots + Y_t - \theta_t}{t-s+1} \le c_{i:k} - \theta_0 \right) \right)$$

$$= P_{\theta=\theta^*}\left(\bigcup_{j=i}^{k} \left(\max_{1 \le s \le j} \min_{j \le t = i} \frac{Y_s - \theta_0 + \ldots + Y_t - \theta_0}{t-s+1} \le c_{i:k} - \theta_0 \right) \right)$$

$$= P\left(\max_{1 \le s \le i} \frac{Y_s + \ldots + Y_i}{i-s+1} \le c_{i:k} \right)$$

$$= P(U_i \le c_{i:k}),$$

where Y_i's are i.i.d with density $g(\cdot;\, 0, \frac{1}{n})$,

$$\theta^* = (\underbrace{\theta_0, \theta_0, \ldots, \theta_0}_{i+1}, \infty, \ldots, \infty)$$

and

$$U_i = \max_{1 \le s \le i} \frac{Y_s + \ldots + Y_i}{i-s+1} .$$

Corollary 1.4.1. $c_{i:k} = c_{i:i}$, $i = 1, \ldots, k$.

For any $x > 0$, let $S_n = \sum_{i=1}^{n} (Y_i - x)$, $n = 1, 2, \ldots$, $S_0 = 0$. Since $0 < P(Y_i - x < 0) = G(x, 0, \frac{1}{n}) < 1$, the distribution of $Y_1 - x$ is not concentrated on a half-axis. By Theorem 1. [illegible] the probability generating function of cumulative distribution functions $P(U_k \le x)$, $x > 0$, $k = 1, 2, \ldots$, is given by

$$\exp\left\{ \sum_{k=1}^{\infty} \frac{s^k}{k} G(xkn, k, 1) \right\} .$$

Hence by Theorem 1.3.6, we have the following recurrence formula for all $x > 0$

$$P(U_{k+1} \le x)$$

$$= \frac{1}{k+1} \sum_{j=0}^{k} P(U_{k-j} \le x) \cdot G(x(j+1)n; (j+1), 1) \quad (1.4.4)$$

where

$$P(U_0 \le x) = 1.$$

When $x = 0$, both sides of Equation (1.4.4) equal to zero, hence it also holds for $x = 0$.

Note that

$$P\left(\frac{1}{r} \sum_{i=1}^{r} Y_i \le x\right) = G\left(x; r, \frac{1}{rn}\right)$$

$$= G(xrn; r, 1). \quad (1.4.5)$$

The values $c_{k:k}(P^*, r, n)$ which satisfy Equation (1.4.5) are tabulated in Table IV for $k = 1 (1) 10$, $P^* = .99, .95, .90, .75$, $r = 2, 4, 6$, and $n = 4, 6, 8, 10, 15, 20$.

Lemma 1.4.1. $c_{i:k} \le c_{i+1:k}$ for all $1 \le i \le k - 1$.

Proof. The constants $c_{i:k}$ ($i = 1,\dots,k$) are determined by (1.4.) respectively.

$U_{i+1} \ge U_i$ a.s. implies $c_{i:k} \le c_{i+1:k}$ for all $1 \le i \le k - 1$.

Theorem 1.4.2. The selection procedure will not be changed if the isotonic estimators $\bar{Y}_{i:k}$, $i = 1,\dots,k$ are replaced by [illegible], $i = 1,\dots,k$

where

$$X_{i:i} = \max_{1 \le s \le i} [illegible]$$

$$= \max_{1 \le s \le i} \frac{X_s + \dots + X_i}{i-s+1}$$

Proof. The proof is similar to that of Theorem 1.[illegible]

Next, we define a selection procedure by using an [illegible] estimator and a fixed constant which depends on P*, k, [illegible] size n and common [illegible].

Definition 1.4.2. The selection procedure R_7 is defined by

R_7: Select π_i if and only if $X_{i:k}$ [illegible] $i = 1, \dots, k$

where c ([illegible] 1) is the smallest value such that procedure [illegible] the P*-condition.

Corollary 1.4.2. The constant $c = c(P^*, k, n)$ of the selection procedure R_7 equals to $c_{k:k}$ which is determined by [illegible]

Proof. Follows immediately from Theorem [illegible] and [illegible]

Definition 1.4.3. The selection procedure [illegible]

Step 1. Select [illegible]

[illegible]

Otherwise, select [illegible]

Step 2. Select π_j, j [illegible] $k-$ [illegible] and stop, if

$$\frac{X_{k-1}}{Y_{k-1}} \geq c_{k-1}(\theta^*)$$

otherwise reject π_{k-1} and go to step 3.

$\vdots$

Step k-1. Select π_j, j [illegible] 2 and stop, if

$$\frac{X_2}{Y_2} \geq c_2(\theta^*)$$

otherwise reject π_2 and go to step k.

Step k. Select π_1 and stop, if

$$\frac{X_1}{Y_1} \geq c_1(\theta^*)$$

otherwise reject π_1.

Here the c_j's are the smallest real values [illegible] such that the procedure [illegible] satisfies the P*-condition.

[illegible] 1.3. [illegible] The constants c_i are determined by

[illegible] $i+1$ [illegible] -1 [illegible] P*, $i = 1, \ldots,$ [illegible]

[illegible] arbitrary i, 1 [illegible]

[illegible]

$$= \inf P\left(\bigcup_{i=1}^{k} \left(\cdots \right)\right)$$

$$= \inf_{i} P\left(\bigcup_{j=i}^{k} \left(\cdots c_i \cdots \theta_i \right)\right)$$

$$= P_{\star}\left(\bigcup_{j=i}^{k} \left(\cdots X_j \cdots c_i \cdots \theta_i^{\star} \right)\right)$$

$$= P(Z_i \cdots c_i \cdots)$$

where Z_i … X_i, i … $1, \ldots, k$ are i.i.d. with the gamma density

$g(\cdot;$ … $, 1)$, … $\star = (\underbrace{\theta, \ldots, \theta}_{i+1}, \ldots)$. Hence c_i, i … $1, \ldots,$ …

determined by (1.4.6). If … $_1$ … $\ldots$ … $_k$, then c_1 … $\ldots$ … $_k$.

The following selection procedure … was given by Gupta and …

(1958).

Definition 1.4.4. The selection procedure … is defined by

…: Select π_i if and only if … c' … i … $1, \ldots, k$

where c' is determined by

$$\prod_{i=1}^{k} \cdots \int_0^{\cdots} \cdots e^{-u} du \cdots P^{\star}.$$

For … $(i = 1, \ldots, k)$, it turns …

$$\cdots \int_0^{\cdots} \cdots du \cdots (P^{\star})^{\frac{1}{k}}$$

the left hand side is the c.d.f. of $\frac{1}{2}\chi^2_\nu$ with ν degrees of freedom, hence the value c' can be easily solved with the help of a table of chi-square distribution.

Application to the Selection of Variance of Normal Population

If $\theta_i = \sigma_i^2$, $i = 0, 1, \ldots, k$ are the scale parameters for the $k+1$ normal populations and X_{ij} ($j = 1, \ldots, n$; $i = 1, \ldots, k$) are the n observations on the population π_i with the mean μ_i (known). We assume that the order $\sigma_1^2 \le \ldots \le \sigma_k^2$ is known.

In the application of selection procedure R_6 or R_7, what we need to do is to evaluate the isotonic regression of S_i^2 which is the sample variance of population π_i, $i = 1, \ldots, k$ and denote it by $S_{i:k}^2$, $i = 1, \ldots, k$, then directly apply R_6 or R_7. The constant we need is determined by Equations (1.4.2) and (1.4.4) where we replace ν by n, the reason being that nS_i^2/σ_i^2 has χ_n^2 distribution with n degrees of freedom and $Y_i = S_i^2/\sigma_i^2$ has the c.d.f.

$$G(2nt;\ n,\ 2) = G(t;\ n,\ \tfrac{1}{n}),$$

hence

$$P\left(\frac{1}{r}\sum_{i=1}^{r} Y_i \le t\right) = G(trn;\ rn,\ 1).$$

The application of R_8 is similar to that of R_4 (see Gupta [illegible] (1982)). What we need to do is to replace $\bar{X}_i$ in R_3 and R_4 by S_i^2 and replace ν in Equation (1.4.6) and (1.4.7) by n, $i = 1, \ldots, k$.

Remark 1.4.1. δ_9 (Gupta and Sobel (1958)) does not depend on the [illegible] ing prior and the sample sizes for each population need not be [illegible]

If the means μ_i, $i = 1,\ldots,k$ are unknown and common sample [illegible] $n > 1$, let $S_i^2 = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2/n-1$ and use $n - 1$ in place of [illegible] (1.4.4), (1.4.6) and (1.4.8) which determined the constants [illegible] c and c' for δ_6, δ_8, and δ_9, respectively.

1.4.2. Selection Procedure $\delta_i^{(2)}$, $i = 6, 7, 8, 9$

Case II. θ_0 unknown.

The assumptions are the same as in Case I except that n_0 [illegible] tions, viz., $x_{01},\ldots,x_{0n_0}$ are taken on π_0.

For selection procedure $\delta_6^{(2)}$, the inequalities defining the pro- cedure and corresponding to $X_{i:k} \geq c_{i:k}\theta_0$ ($i = 1,\ldots,k$) is replaced by $X_{i:k} \geq c_{i:k}^{(2)} Z_0$ ($i = 1,\ldots,k$), respectively. The equation determining $c_{i:k}^{(2)}$ is obtained as before and is given by

$$\int_0^{\infty} P(U_i \geq c_{i:k}^{(2)} t) f(t)\,dt = P^* \quad [illegible]$$

where $P(U_i \ldots)$ is the same as that in Theorem 1.4.1, and $f(t)$ [illegible] p.d.f. of Z_0 of population π_0.

If population π_0 has gamma distribution with density [illegible] (α_0 known and θ_0 unknown), then [illegible] (x; $n_0\alpha_0$, $\frac{1}{n_0}$).

For selection procedure $\delta_7^{(2)}$, the inequality defining the procedure is

$$X_{k:k} \geq c' X_0$$

and it can be shown $c' = c^{(\cdot)}_{k:k}$.

For selection procedure $R_8^{(?)}$, the inequality defining the procedure and corresponding the $(k-i)$th step is

$$\frac{X_i}{\nu_i} \geq c \frac{X_0}{\nu_0} \quad \text{where } \nu_0 = n_0 \theta. \qquad (1.\ldots)$$

The equation determining c of $R_8^{(?)}$ is given by

$$\int_0^\infty \int_0^{ct\,\nu_i/\nu_0} \frac{u^{\frac{\nu_i}{2}-1} e^{-u}}{\Gamma(\frac{\nu_i}{2})}\, du\, \frac{t^{\frac{\nu_0}{2}-1} e^{-t}}{\Gamma(\frac{\nu_0}{2})}\, dt = P^*. \qquad (1.\ldots)$$

For selection procedure $R_9^{(?)}$, the inequality defining the procedure is

$$\frac{X_i}{\nu_i} \geq c \frac{X_0}{\nu_0}, \qquad (1.\ldots)$$

and the equation determining c is given by Gupta and Sobel (19..) as follows:

$$\int_0^\infty \prod_{i=1}^{k} \int_0^{ct\,\nu_i/\nu_0} \frac{u^{\frac{\nu_i}{2}-1} e^{-u}}{\Gamma(\frac{\nu_i}{2})}\, du\, \frac{t^{\frac{\nu_0}{2}-1} e^{-t}}{\Gamma(\frac{\nu_0}{2})}\, dt = P^*. \qquad (1.\ldots)$$

1.5. Selection Rules for the Location Parameter under Partial Ordering Prior Assumption

Assume that we have only a partial ordering prior of k location parameters, that is, the parameter space

[illegible] there is a partial [illegible]

[illegible] to partition [illegible]

sets, say $B_0, \ldots$, so that [illegible]

and for each B_i [illegible] $1, \ldots,$ [illegible]

is no order relation among the elements [illegible]

Let b_i [illegible], the number of elements [illegible]

so we have

$$\sum b_i = [illegible]$$

If we denote the new induced [illegible]

a parameter space [illegible]

an induced partial order.

Example. Suppose k [illegible], and we have a partial ordering [illegible]

[illegible] and [illegible]

represent this [illegible] in Figure 1.

Figure 1. [illegible]

Then we have an induced partial ordering [illegible] as in Figure 2.

Figure 2. Induced partial ordering.

And

B_0 [illegible] $\{5, 8\}$

B_1 [illegible] $\{1, 2, 3, 4\}$

B_2 [illegible] $\{6, 7\}$

It is clear that the induced partial order is not unique, for example, we can partition [illegible] into three other subsets B'_0, B'_1, B'_2 where

B'_0 [illegible] $\{5, 8\}$

B'_1 [illegible] $\{1$, [illegible]$\}$

B'_2 [illegible]

For the location parameter case, a selection procedure [illegible] is defined as follows:

Definition 3.2.1. We define a selection procedure [illegible] as follows:

Suppose $B_0, \ldots, B_s$ are the induced subsets and that for each [illegible] $i = 0, 1, \ldots,$ there is no [illegible] edge on it. We choose a proper [illegible]

selection procedure [illegible] for each subset [illegible] the [illegible]

probability of a correct selection is not less than [illegible]

subset Ω_0 we may use selection procedure [illegible] or [illegible] with [illegible]

Theorem 1.5.1 [illegible] The selection procedure [illegible]

[illegible]

$$\inf P(CS \mid [illegible])$$

$$\inf P(CS \mid R) [illegible]$$

$$\prod_{i=1} \inf_{\Omega_i} [illegible]$$

[illegible]

where Ω_i is the parameter space associated with the subset [illegible]

Remark 1.5.1. For the selection problem of the [illegible] gamma distributions, [illegible] selection procedure can be [illegible]

1.6. Comparisons of the [illegible]

1.6.1. The Location Parameters of Normal Distribution

In this section we use Monte Carlo technique [illegible] formance of selection procedures [illegible] and [illegible]

k independent populations, each population with distribution $N(\theta_i, \sigma^2)$, with common known variance σ^2 and common sample size n. Assume that the mean θ_0 of the control is known; without loss of generality we assume that $\theta_0 = 0$ and $\sigma/\sqrt{n} = 1$.

In the simulation, we use Rubin and Hinkle's RVP-Random Variable Package, Purdue University Computing Center, to generate random numbers. For each k, we generated one random number (variable) for each population, then applied each selection procedure separately and repeated it ten thousand times; we used the relative frequencies as an approximation of the exact values of the associated performance characteristics for each procedure. In Table V we use the following notations:

$\theta = (\theta_1, \ldots, \theta_k)$, θ_i is the parameter of population π_i.

PS = P(CS)

PI = P(correctly rejecting all bad populations)

PC = P(correct classification of all populations)

where the correct classification means that we select all good populations and reject all bad populations.

EI = Expected number (size) of bad populations contained in the selected subset.

$$EL = \sum_{\theta_i > \theta_0} (\theta_i - \theta_0)^2 P(\pi_i \text{ is selected})$$

ES = Expected size of the selected subset.

Table V.1 consists of four parts, namely, the [illegible] $k = 2, 3, 4, 5$. For each value of k we assume that the [illegible] ulation is the one and only one bad population with [illegible] is less than the control [illegible]. A glance at the [illegible] the performance can roughly be ordered as follows:

[illegible]

i.e. procedure [illegible] is the best one and [illegible] and [illegible] are very close and both are better than [illegible] is based on the characteristics (1), (2), (3) [illegible] in Section [illegible]. As the number of populations k [illegible] to five and the three additional populations [illegible] parameter 1, [illegible], and 3, respectively, we find that [illegible] ($i = 1, 2, 3, 4$) is 0.0124, 0.0[illegible], 0.00[illegible], [illegible] This means that when k increases and the additional [illegible] [illegible], then [illegible] is the most sensitive procedure [illegible].

Table V.[illegible] has the [illegible] assume that we have two bad [illegible] on the performance characteristics [illegible] formance is ordered [illegible]

[illegible]

In Table V.[illegible] [illegible]

[illegible]

[illegible]

by [illegible]

populations, then [illegible]

This is because [illegible]

In Table V.4 we assume [illegible] population is bad. [illegible] find that performance is [illegible] the same result as before.

In Table V.5 we assume that the [illegible] is incorrect; i.e. the true [illegible] by [illegible] on P1, P2, [illegible] and [illegible] we have [illegible] again [illegible] is the best. If we [illegible] that [illegible] does not change (the small [illegible] fluctuations), [illegible] and [illegible] increase [illegible] example, in case [illegible] 6, [illegible] increases [illegible] by [illegible] increases [illegible]

From these five tables, [illegible] of these procedures [illegible] correct, otherwise [illegible] than the others.

[illegible] Performance of [illegible]

In this section, we [illegible] the [illegible] procedures [illegible]

populations has exponential distribution [illegible]

[illegible] with unknown scale parameters [illegible]

In the simulation study, [illegible]

we used Rubin and Sickles's PVP [illegible]

variables) with specified exponential distributions. [illegible]

results are based on two thousand sets of random [illegible]

tations in Table VI.[illegible] and Table VI.[illegible] [illegible]

[illegible].6.1 but the meaning of a bad population i[illegible]

a bad population has a parameter which is greater [illegible]

(assumed one).

A good population means that its parameter is less [illegible]

to one. The results of Table VI.1 and Table VI.[illegible] [illegible]

that we have the performance [illegible] [illegible]

that [illegible] is slightly better than [illegible] [illegible] there is [illegible]

[illegible]

Table of $d_{1:k}^{(1)}$ values [illegible] carry out the procedure [illegible] for the normal means problem [illegible] ordering prior.

$d_{1:k}^{(1)}$	.99	.975	[illegible]	[illegible]	[illegible]
[illegible]	2.3264	1.9600	[illegible].6449	1.4[illegible]	1.[illegible]
2	[illegible].33[illegible]7	1.9775	1.6[illegible]50	1.48[illegible]	1.4[illegible]
[illegible]	2.33[illegible]9	1.9787	[illegible].6817	1.49[illegible]	1.[illegible]
4	--	--	[illegible]	1.4[illegible]	[illegible]
5	--	--	1.68[illegible]4	1.49[illegible]	[illegible]
6	--	-	--	--	[illegible]
[illegible]	[illegible].3340	1.9787	1.6[illegible]24	1.4[illegible]	[illegible]

$d_{1:k}^{(1)}$	.8[illegible]	.8[illegible]	[illegible]	.70	[illegible]
1	1.0364	.8416	[illegible]	[illegible]	[illegible]
[illegible]	1.1[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.14[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.14[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.1[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

Table of $d_{1:k}$ values [illegible]

procedure [illegible] for the normal [illegible]

$d_{1:k}$					
[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]50	[illegible]	[illegible]		
	[illegible]	[illegible]	[illegible]		
	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	.4[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible] 761	[illegible]	[illegible]		
[illegible]	[illegible]76	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	1.[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]

TABLE III

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.4420	[illegible]	1.7361	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.4855	1.4378	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.5018	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.6311	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]	1.4637	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.2041	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.4291	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1864	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.0427	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	1.3582	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible] of the procedure [illegible] for the [illegible] multiple observation.

[illegible]

[illegible]

[illegible]

	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	10	--	--	--	--	--	
	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	4.0113	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	--	--	--	--	
	10	[illegible]	[illegible]	--	--	--	
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	--	--	--	--	--	
	[illegible]	--		--	--	--	

[illegible]

[illegible]

	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
	[illegible]	[illegible]	[illegible]	[illegible]			
	[illegible]	–	–	–			
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
	3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	–	–	–			
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]				

[illegible] (continued)

[illegible]

r	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1.67[illegible]	1.5[illegible]	1.[illegible]71[illegible]	1.4[illegible]	[illegible]	[illegible]
[illegible]	1.7[illegible]8	1.5[illegible]	1.56[illegible]	1.4505	[illegible]	[illegible]
[illegible]	1.7406	1.5[illegible]	1.5[illegible]	1.456[illegible]	[illegible]	[illegible]
4	1.7440	1.5996	1.5157	1.4579	[illegible]	[illegible]
5	1.7451	1.6006	1.5156	1.4584	[illegible]	[illegible]
6	1.7455	1.6008	1.5157	1.45[illegible]	[illegible]	[illegible]
7	1.7456	1.6009	1.5159	1.4586	--	--
8	1.7457	--	--	--	--	--
10	--	--	--	--	--	--
1	2.9427	2.7667	2.6615	2.59[illegible]	[illegible]	[illegible]
[illegible]	3.0126	2.8191	2.70[illegible]	2.6[illegible]79	[illegible]	[illegible]
[illegible]	3.0264	2.8233	2.71[illegible]	2.6350	2.5144	[illegible]
4	3.0301	2.8320	2.7155	2.6368	2.5158	[illegible]
5	3.0313	2.8328	2.71[illegible]	2.6374	2.5162	[illegible]
6	3.0317	2.8331	2.71[illegible]	2.6376	2.5164	[illegible]
7	3.0318	2.83[illegible]	2.71[illegible]	2.6377	--	[illegible]
8	3.0319	-	--	--	--	
10	--	--	--	--	-	--
1	4.149[illegible]	[illegible]	[illegible]	3.7[illegible]	3.6266	[illegible]
[illegible]	4.[illegible]	[illegible]68	[illegible]	3.7636	[illegible]	[illegible]
[illegible]	4.[illegible]	4.006[illegible]	3.866[illegible]	3.7717	[illegible]	[illegible]
4	4.[illegible]480	4.0092	3.868[illegible]	3.7737	[illegible]	[illegible]
5	4.249[illegible]	4.0101	3.86[illegible]	3.7744	3.627[illegible]	[illegible]
6	4.249[illegible]	4.01[illegible]4	[illegible]	3.7746	[illegible]	[illegible]
[illegible]	4.2498	[illegible]	[illegible]	3.7747	-	[illegible]
10		-	--	--		

[illegible]

[illegible]

K = [illegible], [illegible]

	1			
PS	.9463	[illegible]	[illegible]	[illegible]
PI	.3864	[illegible]	[illegible]	[illegible]
PR	.3307	[illegible]	.3516	[illegible]
EI	.6146	[illegible]	.6063	[illegible]
ED	.6146	[illegible]	[illegible]	[illegible]
ES	1.5595	[illegible]	1.5642	[illegible]

K = 3, (-1,0,1)

	1			
PS	[illegible]	[illegible]	[illegible]	[illegible]
PI	[illegible]	[illegible]	[illegible]	[illegible]
PR	[illegible]	[illegible]	[illegible]	[illegible]
EI	[illegible]	[illegible]	.6124	[illegible]
ED	[illegible]	[illegible]	.6124	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

K = 4, (-1,0,1, [illegible]

	1			
PS	[illegible]	[illegible]	[illegible]	[illegible]
PI	.3664	[illegible]	[illegible]	[illegible]
PR	[illegible]	[illegible]	[illegible]	[illegible]
EI	.6336	[illegible]	[illegible]	[illegible]
ED	.6336	[illegible]	[illegible]	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

K = [illegible], (-1,0,1, [illegible]

	1			
PS	[illegible]	[illegible]	[illegible]	[illegible]
PI	[illegible]	[illegible]	[illegible]	[illegible]
PR	[illegible]	[illegible]	[illegible]	[illegible]
EI	[illegible]	[illegible]	[illegible]	[illegible]
ED	[illegible]	[illegible]	[illegible]	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

Simulation results [illegible]

[illegible]

S	1. [illegible]	[illegible]	[illegible]	[illegible]
PI	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	.[illegible]7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

PS	[illegible]	[illegible]	[illegible]	[illegible]
PI	[illegible]7	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

PS	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		[illegible]
[illegible]	[illegible]		[illegible]
[illegible]	[illegible]		[illegible]
[illegible]	[illegible]		[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

Simulation results for the comparative performance of [illegible] selection procedures for the normal means [illegible] (notation explained [illegible]) under simple ordering [illegible].

p* [illegible]

k [illegible]

	1			
PS	1.0000	[illegible]	[illegible]	[illegible]
PI	.9613	[illegible]	[illegible]	[illegible]
PC	.9613	[illegible]	[illegible]	[illegible]
EI	.0387	[illegible]	[illegible]	[illegible]
EO	.3563	[illegible]	[illegible]	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

k [illegible]

	1			
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

k [illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	.9932	[illegible]
[illegible]	.3376	[illegible]	[illegible]	[illegible]
[illegible]	.6080	[illegible]	[illegible]	[illegible]
[illegible]	1.7915	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

k [illegible]

	1			
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	.1112	[illegible]	[illegible]	[illegible]
[illegible]	.1112	[illegible]	[illegible]	[illegible]
[illegible]	.7800	[illegible]	[illegible]	[illegible]
[illegible]	1.4115	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE V. [illegible]

Simulation results for the comparative performance of [illegible] selection procedures for the normal means problem (notation explained in Section 1.6.1) under simple ordering [illegible].

$P^* = .900$

k = 2, [illegible] = (-1,-2)

	1	[illegible]	[illegible]	[illegible]
PS	1.0000	1.0000	1.[illegible]	1.[illegible]
PI	.5405	.534[illegible]	.[illegible]937	.[illegible]72[illegible]
PC	.5405	.5349	.[illegible]937	.[illegible]
EI	.8331	.8[illegible]87	1.3151	1.[illegible]04
EJ	[illegible].2116	[illegible].2[illegible]40	[illegible].4[illegible]32	[illegible]
ES	.8331	.8[illegible]87	1.3151	1.0904

k = 3, [illegible] = (-1,-[illegible],1)

	1	[illegible]	[illegible]	[illegible]
PS	.995[illegible]	[illegible]	.9967	.[illegible]
PI	.5365	.5[illegible]	.[illegible]7	.[illegible]19[illegible]
PC	.5297	[illegible]	.[illegible]944	.[illegible]
EI	.8347	.[illegible]6[illegible]	1.[illegible]116	[illegible]15[illegible]
EJ	[illegible]	[illegible].2[illegible]1[illegible]	[illegible].4156	[illegible]19
ES	1.8279	[illegible].8[illegible]	[illegible]07[illegible]	[illegible]

k = 4, [illegible] = (-1,-2,1,[illegible])

	1	[illegible]	3	[illegible]
PS	.9921	.9[illegible]	.[illegible]7[illegible]	.[illegible]6
PI	.5271	.5[illegible]	.[illegible]44	.[illegible]
PC	.519[illegible]	.5[illegible]	.[illegible]67	[illegible]
EI	.8498	[illegible]	1.[illegible]	[illegible]77
EJ	2.2685	[illegible]	[illegible]	.[illegible]7[illegible]
ES	[illegible].839[illegible]	[illegible]	[illegible]	[illegible]

k = [illegible], [illegible] = (-1,-2,1,[illegible]

	1	[illegible]	[illegible]	[illegible]
PS	.[illegible]	[illegible]	[illegible]	[illegible]
PI	.5317	[illegible]	[illegible]	[illegible]
PC	[illegible]	[illegible]	[illegible]	[illegible]
EI	.[illegible]461	[illegible]	[illegible]	[illegible]
EJ	[illegible]51	[illegible]	[illegible]	[illegible]
ES	[illegible]841	[illegible]	[illegible]	[illegible]

TABLE [illegible]

Simulation results for the comparative [illegible] procedures for the gamma means problem [illegible] 1.6.1) under simple ordering prior.

P^{*} [illegible] .500

k [illegible] 4, [illegible] (.3,.6,.9,1.[illegible])

PS	.9840	[illegible]	[illegible]	[illegible]
PI	.2020	[illegible]	[illegible]	[illegible]
PC	.1860	[illegible]	[illegible]	[illegible]
EI	.7980	[illegible]	[illegible]	[illegible]
EJ	.0319	[illegible]	[illegible]	[illegible]
ES	[illegible]	3.8190	[illegible]	[illegible]

k [illegible] (.[illegible],.6,.9,1.[illegible],1.5)

PS	.9905	.9910	.9965	[illegible]
PI	.0990	.0950	.0645	[illegible]
PC	.0895	[illegible]	[illegible]	[illegible]
EI	1.5040	[illegible]	[illegible]	[illegible]
EJ	.1868	[illegible]	[illegible]	[illegible]
ES	4.491[illegible]	[illegible]	4.5[illegible]	[illegible]

k [illegible] 4, [illegible] (.[illegible],1,1.[illegible],[illegible])

PS	.9685	[illegible]	[illegible]	[illegible]
PI	.2685	[illegible]	[illegible]	[illegible]
PC	.2370	[illegible]	[illegible]	[illegible]
EI	1.0790	[illegible]	[illegible]	[illegible]
EJ	.5304	[illegible]	[illegible]	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

k [illegible] (.[illegible],1,1.[illegible],[illegible])

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE VI.

Simulation results for the comparative performance of various selection procedures for the gamma means problem (notation explained in Section 1.6.1) under simple ordering prior.

$P^* = .900$

k = 4, (.1,.6,1.1,1.6)

	6	7	8	[illegible]
PS	.9995	.9935	1.0000	.9995
PI	.0690	.0645	.0545	.0110
PC	.0685	.0640	.0545	.0110
EI	1.4940	1.5345	1.5315	1.7500
EJ	.2120	.2250	.2204	.2940
ES	3.4935	3.5340	3.5315	3.7495

k = 5, (.1,.6,1.1,1.6,2.1)

	6	7	8	[illegible]
PS	.9990	.9990	1.0000	.9995
PI	.0540	.0530	.0295	[illegible]
PC	.0530	.0520	.0295	[illegible]
EI	1.9205	1.9700	2.1385	[illegible]
EJ	.6140	.6646	.7727	[illegible]
ES	3.9195	3.9630	4.1385	[illegible]

k = 4, (1.1,1.4,1.7,1.9)

	6	7	8	[illegible]
PS	1.0000	[illegible]	1.0000	[illegible]
PI	.0375	[illegible]	.0095	[illegible]
PC	.0375	[illegible]	.0085	[illegible]
EI	.6890	[illegible]	[illegible]	[illegible]
EJ	.6796	[illegible]	[illegible]	[illegible]
ES	2.6890	[illegible]	[illegible]	[illegible]

k = 5, (1.1,1.4,1.7,[illegible])

	6	[illegible]	[illegible]	[illegible]
PS	1.0000	[illegible]	[illegible]	[illegible]
PI	.0315	[illegible]	[illegible]	[illegible]
PC	.0315	[illegible]	.0050	[illegible]
EI	[illegible]	[illegible]	[illegible]	[illegible]
EJ	[illegible]	[illegible]	[illegible]	[illegible]
ES	[illegible]	[illegible]	[illegible]	[illegible]

CHAPTER II

BAYES - P* SELECTION RULES

FOR SELECTING A SUBSET CONTAINING

THE BEST POPULATION

2.1. Introduction

Suppose we have k(k [illegible] 2) independent populations [illegible] the random variable X_i associated with [illegible] has a distribution with [illegible] known parameters [illegible], i = 1, ... [illegible]. First, we give some definitions.

Definition 2.1.1. The populaiton π_i is the best population [illegible] for all [illegible] ≠ [illegible]. If there are more than one populations [illegible] condition we arbitrarily tag one of them and call it the [illegible] population which is not the "best" is called a non-best [illegible]

Assume that we have n_i independent observations [illegible] for population [illegible], i = 1,...,k [illegible] estimator [illegible], i = 1, ..., k, so that [illegible] are independent [illegible] for each [illegible], we have [illegible]

[illegible]

[illegible]

$\{1,2,\ldots,k\}$. An action $A \in \mathcal{Q}$ is the selection of some subset of the k populations. $i \in A$ means that π_i is included in the selected subset. An action $A \in \mathcal{Q}$ is called a correct selection (CS) if the best population is included in the selected subset A.

Definition 2.1.2. A measurable function δ defined on [illegible] is called a selection procedure provided that for each $x \in \mathcal{X}$, we have

$$0 \le \delta(x,A)$$

and

$$\sum_{A \in \mathcal{Q}} \delta(x,A) = 1$$

where $\delta(x,A)$ dentoes the probability that the subset A is selected when x is observed.

The individual selection probability $\psi_i(x)$ for the population [illegible] is then given by

$$\psi_i(x) = \sum_{A \ni i} \delta(x,A)$$

where the summation is over all subsets A which contain i. [illegible] lection probability $\psi_i(x)$ takes on only value 0 or 1, [illegible] the selection procedure $\delta(x,A)$ is completely specified [illegible] is called a non-randomized procedure.

Definition 2.1.3. Two selection procedures [illegible] and [illegible] are equivalent if they have the same individual selection probabilities [illegible] for all x, $i = 1,\ldots,k$.

Hence we can use the following definition [illegible]

Definition 2.1.4. A subset selection rule δ is a measurable [illegible] from $\mathfrak{X}$ to $\mathbb{R}^k$,

$$\delta(x) = (\delta_1(x), \ldots, \delta_k(x))$$

with

$$0 \le \delta_i(x) \le 1, \quad i = 1, \ldots, k.$$

If δ_i's are 0 or 1, the rule is non-randomized. [illegible] rition 2.1.[illegible], we have $\sum_{i=1}^{k} \delta_i(x) \ge 1$.

Suppose our goal is to find a nontrivial subset which [illegible] best population. A large body of literature exists in the area [illegible] set selection procedures (see Gupta and Panchapakesan ([illegible]) [illegible] (1956, 1965) gave maximum-type subset selection procedure. [illegible] Hsu (1979) studied the performance of Gupta-type maximum procedure, Seal-type average procedure (Seal (1955, 1957)) and the [illegible] Berger (1979) and Berger and Gupta (1980) proved that Gupta-type [illegible] procedure [illegible] minimax under certain loss function. In the [illegible] retic approach to the subset selection problem, [illegible] Chernoff and Yahav (1977), Bickel and Yahav (1977), Gupta and [illegible] and Kim (1979) gave different formulation under different [illegible] The loss functions proposed [illegible] consist of two [illegible] [illegible] and the results [illegible] to the ratio of the [illegible] ents of the two components. Hence, whenever the ratio [illegible] may wish to try some other method of attack.

On the other hand, as pointed out by [illegible] Ramal[illegible] ([illegible]) the [illegible] of homogeneity of [illegible] [illegible]

[illegible] [illegible] [illegible]

[illegible] however, [illegible]

[illegible], usually [illegible] when the distributions are [illegible]

der the hypothesis of homogeneity [illegible]

[illegible] of homogeneity is rejected at some small significance level [illegible]

we may wish to relax (modify) the so-called P^*-condition.

In this chapter we define the posterior-P^* condition [illegible] Two Bayes-P^* selection procedures [illegible] and [illegible] are proposed [illegible] and Section 2.3 separately, and their properties are [illegible] 2.4. In Section 2.5 we discuss their applications to [illegible] tions. In Section 2.6 procedure [illegible] is compared with [illegible] procedure. An application for the problem of selecting [illegible] $i = 1,\ldots,k$ of the normal distributions $N(\mu_i, \sigma_i^2)$ [illegible] in Section 2.7. In Section [illegible] we discuss their [illegible] selection problems for Poisson distributions and [illegible] their relation to the selection of gamma distributions [illegible] deals with comparisons of the performance of [illegible] and [illegible] where [illegible] and [illegible] are the [illegible] cedures based on [illegible] and sample [illegible] Gupta (1956, 1965) and Gupta and Singh [illegible] [illegible] on Monte Carlo studies [illegible] [illegible] in terms of the expected selected size [illegible] [illegible] [illegible] [illegible]

2.2. Definitions of the Posterior-P^* Condition and the [illegible] Bayes-P^* Procedure δ_{NR}^{B}

Let $\theta_{[1]} \le \dots \le \theta_{[k]}$ be the ordered unknown θ_i's. [illegible] prior distribution [illegible] for $\theta = (\theta_1, \dots, \theta_k)$ then the posterior [illegible] of a correct selection under selection procedure δ, given [illegible]

$$P(CS|\delta, X = x) = \sum_{i=1}^{k} \delta_i(x) p_i(x)$$

where

$$p_i(x) = P(\pi_i \text{ is the best} | X = x).$$

It is clear

$$\sum_{i=1}^{k} p_i(x) = 1.$$

Definition 2.2.1. Given a number $P^* (\frac{1}{k} < P^* < 1)$ and the prior [illegible], we [illegible] a selection procedure δ satisfies the posterior-P^* condition [illegible]

$$P(CS|\delta, X = x) \ge P^* \quad \text{for all } x.$$

Remark 2.2.1. The posterior-P^* condition is based on the prior distribution [illegible] and is different from the usual so-called P^*-condition.

Definition 2.2.2. The loss function [illegible] is defined by [illegible] where $|A|$ is the size (number) of populations associated with the [illegible] of A. The loss function [illegible] is defined by [illegible]

which is the number of the non-best populations selected by [illegible]

Note that the indicator function

[illegible]

0, otherwise

Definition 2.2.3. Given a number P^* $(\frac{1}{k} < P^* < 1)$ and the prior τ, we define the class $\mathcal{D}_{NR}(\tau, P^*)$ as follows.

$$\mathcal{D}_{NR}(\tau, P^*) = \{\delta \mid \delta \text{ is any non-randomized rule which satisfies the posterior-}P^* \text{ condition}\}.$$

For the sake of convenience sometimes we will use $\mathcal{D}_{NR}$ instead of $\mathcal{D}_{NR}(\tau, P^*)$.

Definition 2.2.4. Given a number P^* $(\frac{1}{k} < P^* < 1)$, a prior τ, and the loss function L, a selection procedure $\delta \in \mathcal{D}_{NR}(\tau, P^*)$ is called a non-randomized Bayes-P^* procedure (rule) if δ is a Bayes rule in the class $\mathcal{D}_{NR}(\tau, P^*)$.

Let $p_{[1]}(x) \leq \ldots \leq p_{[k]}(x)$ be the ordered $p_i(x)$'s and $\pi_{(i)}$ be the population associated with $p_{[i]}(x)$, $i = 1, \ldots, k$, then a subset selection rule δ is completely specified by $\delta_{(1)}, \ldots, \delta_{(k)}$, where $\delta_{(i)}$ is defined by

$$\delta_{(i)}(x) = P(\pi_{(i)} \text{ is selected} \mid X = x), \quad i = 1, \ldots, k.$$

Next, we propose a *non-randomized* selection rule which belongs to $\mathcal{D}_{NR}(\tau, P^*)$.

Definition 2.2.5. Given a number P^* $(\frac{1}{k} < P^* < 1)$, $x \in$ [illegible], and a prior distribution τ, the selection rule δ^{*}_{NR} is defined by [illegible], where

$$\delta_{(i)}(x) = \begin{cases} 1, & \text{if } i \geq [illegible] \\ 0, & \text{otherwise} \end{cases}$$

and $j(x)$ is the maximum integer such that

$$\sum_{i=j}^{k} p_{[i]}(x) \ge P^* .$$

Lemma 2.2.1. $\delta^B_{NR} \in \mathcal{D}_{NR}$.

Proof. Follows from the definition of δ^B_{NR}.

Theorem 2.2.1. Given a number P^* ($\frac{1}{k} < P^* < 1$), the prior τ, and the loss function L_1, the selection procedure δ^B_{NR} is a non-randomized Bayes rule.

Proof. It is sufficient to show that the selection procedure δ^B_{NR} has the smallest posterior risk in the class $\mathcal{D}_{NR}(\tau,P^*)$. Given the observation $X = x$. Let the posterior risk of $\delta \in \mathcal{D}_{NR}(\tau,P^*)$ be $r(x,\delta)$. Then

$$r(x,\delta^B_{NR}) = k - j + 1$$

and

$$\sum_{i=j+1}^{k} p_{[i]}(x) < P^*$$

for some j, $1 \le j \le k$.

Hence the inequality

$$r(x,\delta) < r(x,\delta^B_{NR})$$

is not true for any $\delta \in \mathcal{D}_{NR}(\tau,P^*)$, $\delta \ne \delta^B_{NR}$. Therefore, the result follows.

Theorem 2.2.2. Theorem 2.2.1 also holds when we replace the loss L_1 by L_2.

Proof. Under the loss function L_2, the posterior risk of the selection procedure $\psi \in \mathcal{L}_{NR}(\pi, P^*)$ is

$$r(x,\psi) = \sum_{i=1}^{k} \psi_{(i)}(x)[1 - p_{[i]}(x)], \quad \text{given } X = x.$$

By Theorem 2.2.1, we have

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(x) \le \sum_{i=1}^{k} \psi_{(i)}(x) .$$

If

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(x) = \sum_{i=1}^{k} \psi_{(i)}(x)$$

then by definition of ψ^B_{NR}, we have

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(x)p_{[i]}(x) \ge \sum_{i=1}^{k} \psi_{(i)}(x)p_{[i]}(x) .$$

On the other hand, if

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(x) < \sum_{i=1}^{k} \psi_{(i)}(x)$$

then

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(x) \le \sum_{i=1}^{k} \psi_{(i)}(x) - 1$$

$$\le \sum_{i=1}^{k} \psi_{(i)}(x)(1 - p_{[i]}(x)) .$$

Therefore, we have

$$r(x,\psi^B_{NR}) \le r(x,\psi) \quad \text{for all } \psi \in \mathcal{L}_{NR}(\pi,P^*) .$$

Corollary 2.2.1. For a given prior π and the loss function $L = c_1L_1 + c_2L_2$, where $c_1, c_2 \ge 0$ then ψ^B_{NR} is a non-randomized Bayes

rule wrt the loss function L for all [illegible]

Proof. For the given prior [illegible] and the loss function L, [illegible] risk of any procedure [illegible] is, given [illegible]

$$[illegible] = C_1 \sum_{i=1}^{k} [illegible](x) + C_2 \sum_{i=1}^{k} [illegible]$$

$$= C_1 \sum_{i=1}^{k} [illegible] \sum_{i=1}^{k} [illegible]$$

$$= [illegible](x, [illegible]_{NR}^{B})$$

wrt the loss function L.

Hence $[illegible]_{NR}^{B}$ is a Bayes-P^* rule wrt the loss function L for [illegible]

2.3. Proposed Bayes-P^* Procedure [illegible] in General

Suppose we are interested in the randomized subset selection rule, and we would like to find such a rule which also [illegible] the [illegible] P^* condition and has the minimum risk wrt the loss function [illegible] and the prior distribution [illegible].

Definition 2.3.1. Given a prior [illegible], we define a class [illegible] of selection rules, in which all rules satisfy the [illegible] P^* condition, for any given observation X [illegible], that is,

$$[illegible](\cdot, P^*) [illegible] P^* \text{ for all } [illegible]$$

Definition 2.3.2. Given a number P^* [illegible], a prior [illegible] and loss function L, a selection procedure [illegible] is called a Bayes-P^* [illegible] if this procedure [illegible] is a Bayes rule in the class [illegible].

For the sake of convenience, sometimes we will use [illegible] instead of [illegible]$(\ , P^*)$.

Definition 2.3.2. We define a subset selection procedure $[illegible]^B$ as follows.

Given a prior [illegible] and observation X [illegible] x, $[illegible]^B$ is defined by

$$([illegible]^B_{(1)}, [illegible]^B_{(2)}, \ldots, [illegible]^B_{(k)}),$$

where

$$[illegible]^B_{(k)}(x) = 1$$

and

$$[illegible]^B_{(j)}(x) = 1, \quad \text{if} \quad \sum_{i=j}^{k} p_{[i]}(x) [illegible] P^*, \ j \neq k$$

$$= [illegible], \quad \text{if} \begin{cases} \sum_{k-j+1}^{k} p_{[i]}(x) [illegible] P^* \\ \sum_{i=j}^{k} p_{[i]}(x) [illegible] P^* \\ p_{[j]}(x) + \sum_{i=j+1}^{k} p_{[i]}(x) [illegible] P^* \end{cases}$$

$$0, \quad \text{otherwise.}$$

Example. If $k = 3$, P^* .90 and the posterior probabilities [illegible] $p_{[illegible]}(x)$.[illegible], $p_{[illegible]}(x)$ .15, then we select the population [illegible] corresponding to $p_{[3]}(x)$ with probability 1. And we select [illegible] with probability [illegible] where [illegible] is given by

[illegible]

By Definition 2.3.2 we have

$$[illegible] \overset{B}{,} [illegible] \sum_{i=1}^{k} [illegible]_{(i)}(x) [illegible] P^{*} [illegible]$$

Hence we have the following lemma.

Lemma [illegible] $[illegible]^{B} \in [illegible](\,,P^{*})$.

Definition [illegible]. We define a subclass $[illegible]'(\,,P^{*})$ of class [illegible]

$$[illegible]'(\,,P^{*}) [illegible] \in [illegible](\,,P^{*}) [illegible]_{(i)}(x) [illegible]_{[i]}(x) [illegible]$$

where $[illegible]_{[1]}(x) \ldots [illegible]_{[k]}(x)$ are the ordered $[illegible]_{i}(x)$'s.

By the definition of $[illegible]'(\,,P^{*})$ we have the following lemma.

Lemma 2.3.2. $[illegible]^{B} \in [illegible]'(\,,P^{*})$.

Lemma 2.3.3. For all $[illegible] \in [illegible](\,,P^{*})$ there exists $[illegible]' \in [illegible]'(\,,P^{*})$ [illegible] $(x, [illegible]') [illegible] (x, [illegible])$ wrt the loss function L_1, for all x.

Theorem 2.3.1. Selection procedure $[illegible]^{B}$ is a Bayes-P^{*} procedure in $[illegible](\,,P^{*})$ wrt the loss function L_1.

Proof. Given the observation [illegible] and any selection procedure $[illegible] \in [illegible](\,,P^{*})$, if $[illegible]_{[k]}(x) [illegible] P^{*}$, then [illegible] and [illegible] for $i \neq k$. Hence we have

$$[illegible]^{B} [illegible] \text{ for all } [illegible]$$

If $p_{[k]}(x) [illegible]$, then we have [illegible] and $\sum^{k} [illegible]$

We will now show that for any [illegible]

$$[illegible]^{B} [illegible]$$

implies

$$P(CS|\delta,x) \geq P(CS|\delta^B,x) \geq P^* .$$

That is

$$\sum_{i=1}^{k} \delta_i(x) \leq \sum_{i=1}^{k} \delta_i^B(x)$$

implies

$$\sum_{i=1}^{k} \delta_{(i)}(x) p_{[i]}(x) \leq \sum_{i=1}^{k} \delta_{(i)}^B(x) p_{[i]}(x) .$$

For any C, $1 \leq C \leq \sum_{i=1}^{k} \delta_i^B(x)$, we have $C = a + \gamma$ where a is a positive integer and $0 \leq \gamma < 1$.

It is easy to see that the maximum posterior probability of correct selection of procedure δ with $\sum_{i=1}^{k} \delta_i(x) = C$ is

$$\sum_{i=k-a+1}^{k} p_{[i]}(x) + \gamma p_{[k-a]}(x) .$$

and it is less than $\sum_{i=1}^{k} \delta_{(i)}^B(x) p_{[i]}(x)$, since $C \leq \sum_{i=1}^{k} \delta_{(i)}^B(x)$. Therefore δ^B is Bayes-P^* procedure in [illegible] P^*.

Lemma [illegible].3. Given the loss function [illegible], for all [illegible] P^*, there exists [illegible] $(\cdot, P^*)$ such that, [illegible] is the posterior risk [illegible]

Proof. Given [illegible], let [illegible] that

[illegible]

$$\sum_{i=1}^{k} [illegible]_{(i)}(x) p_{[i]}(x) [illegible]$$

$$[illegible] P^{*}$$

hence [illegible] $\in$ [illegible] $(\;, P^{*})$.

Now,

$$[illegible](x, [illegible]) [illegible] \sum_{i=1}^{k} [illegible](x)(1 - [illegible]_{[i]}(x))$$

$$[illegible] \sum_{i=1}^{k} [illegible]_{[i]}(x)(1 - p_{[i]}(x))$$

$$[illegible] \sum_{i=1}^{k} [illegible]_{(i)}(x)(1 - p_{[i]}(x))$$

$$[illegible](x, [illegible]).$$

Hence the proof is complete.

Theorem 2.3.2. Given the prior [illegible] and the observation [illegible], the procedure [illegible] is a Bayes-P^{*} procedure in the class [illegible] $(\;, P^{*})$ when the [illegible] function is $L_{[illegible]}$.

Proof. By Lemma 2.2.4, it is sufficient to show that

$$[illegible](x, [illegible]) [illegible] \min [illegible]$$

where

$$[illegible](x, [illegible]) [illegible] \sum_{i=1}^{k} [illegible]$$

$$[illegible] \sum_{i=1}^{k} [illegible] \sum_{i=1}^{k} [illegible] p_{[illegible]}(x)$$

[illegible]

$$i_0 = \min_{1 \le i \le k} \{ i : \ldots^{B}_{(i)}(x) \ldots \}$$

$$A_1 = \{ i : \ldots_{(i)}(x) - \ldots^{B}_{(i)}(x) \ldots \}$$

$$A_2 = \{ i : \ldots_{(i)}(x) - \ldots^{B}_{(i)}(x) \ldots \},$$

then $A_1 \cup A_2 = \ldots$ [illegible].

And we have

$$a_1 = \max A_1 = i_0 - 1, \quad \text{if} \quad \ldots^{B}_{(i_0)}(x) \ldots$$

$$= i_0 \quad, \quad \text{if} \quad \ldots^{B}_{(i_0)}(x) \ldots$$

$$a_2 = \min A_2 = i_0 \quad, \quad \text{if} \quad \ldots^{B}_{(i_0)}(x) \ldots$$

$$= i_0 + 1, \quad \text{if} \quad \ldots^{B}_{(i_0)}(x) \ldots$$

hence $a_1 \ldots a_2$.

Therefore, we have

[illegible]

[illegible]

$\cdot (1 - [illegible])$ [illegible]

[illegible] $[\ldots_{(i)}(x) - \ldots^{B}_{(i)}(x)][1 - \ldots]$ [illegible]

[illegible] by Theorem 2.3.1.

Corollary 2.3.1. Procedure [illegible] is a Bayes-P^* rule [illegible] loss function $L = c_1 L_1 + c_2 L_2$, c_1, c_2 [illegible].

Proof. Similar to Corollary 2.2.1, hence it is omitted.

2.4. Properties of [illegible]^{B} and [illegible]$^{B}_{NR}$

In this section we discuss some properties of selection [illegible] [illegible]^{B} and [illegible]$^{B}_{NR}$. The following definition of the ordering of [illegible] was introduced by Lehmann (1955) and further discussed by [illegible] and Alam (197[illegible]).

Definition 2.4.1. A subset A [illegible] R^k is monotone [illegible] z_i [illegible] y_i for all $i = 1, \ldots, k$, implies [illegible].

Definition 2.4.[illegible] A family of [illegible] distributions on R^k, [illegible] R^k [illegible] has the [illegible] property [illegible] and [illegible] for all $i = 1, \ldots, k$, implies

$$[illegible] \int dF [illegible]$$

for all monotone sets [illegible].

Let $f(\cdot, \theta_i)$ be the p.d.f. of population π_i. Let [illegible] be the [illegible] prior where θ_i's are mutually independent. Suppose for [illegible] x, we have absolutely continuous posterior c.d.f. $G(\cdot|x)$. Hence we can write the p.d.f. as

$$g(\theta|x) = \prod_{i=1}^{k} g_i(\theta_i|x) = \prod_{i=1}^{k} g_i(\theta_i|x_i).$$

Let $G_i(\cdot|x_i)$ be the posterior c.d.f. associated with [illegible], $i = 1,\ldots,k$.

Definition 2.4.3. The absolutely continuous posterior c.d.f. $G_i(\cdot|x_i)$, $i = 1,\ldots,k$, have the generalized (strictly) stochastic increasing property, [illegible]G(S)SIP) if for any i, j, $1 \le i$, $j \le k$, [illegible] then

$$G_i(\cdot|x_i) \ [illegible] \ G_j(\cdot|x_j).$$

Note that if [illegible] $G(\cdot|\cdot)$ for all [illegible] then the GSIP is the usual SIP.

Definition 2.4.4. A selection procedure [illegible] is monotone [illegible] only if for every $x \in \mathbb{R}^k$, [illegible] implies [illegible] a.e. if [illegible] is monotone with the exception of [illegible] having probability zero.

Theorem 2.4.1. If the [illegible] that [illegible] independent posterior [illegible] GSIP, then for every $x \in \mathbb{R}^k$, [illegible]

[illegible]

[illegible]

Proof.

$p_i(x)$ [illegible] $P($[illegible]$_i$ [illegible]$_{j \neq j}$ [illegible] $x)$ [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Since

$B_i(x)$ [illegible] $B_j(x)$ [illegible]

and

$B_{NR\,i}(x)$ [illegible] $B_{NR\,j}(x)$ [illegible]

Therefore, the procedures [illegible] and [illegible] are ordered [illegible]

Under [illegible] assumptions, we can relabel the population [illegible] [illegible] hence we have [illegible] Gupta (1970) defined the "just" property of a selection rule [illegible]

Definition 2.1.6. A selection rule [illegible] is called "just" [illegible] for all [illegible] [illegible]

We call [illegible] almost everywhere [illegible] [illegible] probability [illegible]

Definition 2.4.6. A selection rule [illegible] is called translation-invariant [illegible] for all $x \in \mathbb{R}^k$, and for all $c \in$ [illegible]

$$\delta_i(x_1 + c, \ldots, x_k + c) = \delta_i(x_1, \ldots, x_k), \quad i = 1, \ldots, k.$$

Definition 2.4.7. A selection rule [illegible] is called scale-invariant if for all $x \in \mathbb{R}^k$, and for all $c > 0$

$$\delta_i(x_1 \cdot c, \ldots, x_k \cdot c) = \delta_i(x_1, \ldots, x_k), \quad i = 1, \ldots, k.$$

Theorem 2.4.2. If the posterior distributions $G_i(\cdot \mid \cdot)$, $i = 1, \ldots, k$, have the [illegible] property, then both selection procedures [illegible] and [illegible] are "just" i.e. [illegible]

Proof. It is sufficient to show that

$$p_i(x) \leq p_i(x') \quad \text{whenever } x_i' \geq x_i$$

[illegible]

for any fixed i,

[illegible]

[illegible]

AD-A100 948 PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS F/G 12/1

SOME RESULTS IN THE THEORY OF SUBSET SELECTION PROCEDURES.(U)

AUG 80 H YANG N00014-75-C-0455

UNCLASSIFIED MMS-80-22 NL

2 OF 2

AD A
100948

END
DATE
FILMED
7-81
DTIC

Definition 2.4.8. Given a number $P^*(\frac{1}{k} < P^* < 1)$, $\underline{X} = \underline{x}$ and a prior τ; for any selection procedure $\psi \in \mathcal{S}(\tau,P^*)$ the ratio of the posterior probability $P(CS|\psi,\underline{x})$ and the posterior expected selected size $E(S|\psi,\underline{x})$ is called the posterior-efficiency of ψ and is denoted by $EFF(\psi|\underline{x})$.

$$EFF(\psi|\underline{x}) = \frac{P(CS|\psi,\underline{x})}{E(S|\psi,\underline{x})} .$$

If $EFF(\psi|\underline{x}) \geq EFF(\psi'|\underline{x})$ for all $\psi' \in \mathcal{S}$ and all $\underline{x}$, then the selection procedure ψ is called "posterior most efficient" (PME) selection procedure in $\mathcal{S}(\tau,P^*)$.

Theorem 2.4.3. The non-randomized posterior-P^* selection procedure ψ^B_{NR} is the PME selection procedure in $\mathcal{S}_{NR}(\tau,P^*) = \mathcal{S}_{NR}$, given τ, P^*.

Proof. By Lemma 2.2.1, for all $\psi \in \mathcal{S}_{NR}$

$$\exists\, \psi' \in \mathcal{S}'_{NR} \ni EFF(\psi'|\underline{x}) \geq EFF(\psi|\underline{x}) \quad \forall\, \underline{x} ,$$

hence it is sufficient to show that:

Given $\tau(\underline{\theta})$, P^*, $\underline{x}$, $EFF(\psi^B_{NR}|\underline{x}) \geq EFF(\psi'|\underline{x})$ for all $\psi' \in \mathcal{S}'_{NR}(\tau,P^*)$.

We know that, in $\mathcal{S}_{NR}(\tau,P^*)$ hence in $\mathcal{S}'_{NR}(\tau,P^*)$, ψ^B_{NR} always has minimum selected size, i.e. $\forall\, \underline{x}$, $\sum_{i=1}^{k} \psi^B_{NRi}(\underline{x}) + c = \sum_{i=1}^{k} \psi'_i(\underline{x})$ for some integer c, $0 \leq c \leq k-1$.

$$EFF(\psi'|\underline{x}) = \frac{\sum_{i=1}^{k} \psi'_{(i)}(\underline{x})p_{[i]}(\underline{x})}{\sum_{i=1}^{k} \psi'_{(i)}(\underline{x})}$$

$$\leq \frac{\sum_{i=1}^{k} \psi^B_{NR(i)}(\underline{x})p_{[i]}(\underline{x})+p_{[k-s-c+1]}(\underline{x})+\dots+p_{[k-s]}(\underline{x})}{\sum_{i=1}^{k} \psi^B_{(i)}+c}$$

if $\psi^B_{NR}(x) = (0,\ldots,0,\underbrace{1,\ldots,1}_{s \text{ terms}})$.

$$EFF(\psi'|\underline{x}) \le \frac{\sum_{i=k-s+1}^{k} \psi^B_{NR(i)}(\underline{x})p_{[i]}(\underline{x}) + cp_{[k-s]}(\underline{x})}{\sum_{i=1}^{k} \psi^B_{NR(i)}(\underline{x}) + c}$$

$$\le \frac{\sum_{i=1}^{k} \psi^B_{NR(i)}(\underline{x})p_{[i]}(\underline{x})}{\sum_{i=1}^{k} \psi^B_{NR(i)}(\underline{x})}$$

$$= EFF(\psi^B_{NR}|\underline{x}) \quad .$$

The last inequality is obtained by

$$\sum_{i=1}^{k} \psi^B_{NR(i)}(\underline{x})p_{[i]}(\underline{x}) = \sum_{i=k-s+1}^{k} \psi^B_{NR(i)}(\underline{x})p_{[i]}(\underline{x})$$

$$\ge \left(\sum_{i=k-s+1}^{k} \psi^B_{NR(i)}(\underline{x})\right) p_{[k-s]}(\underline{x}).$$

Theorem 2.4.4. The randomized selection procedure ψ^B is the PME procedure in $\mathcal{D}(\tau,P^*) = \mathcal{D}$ for given τ, P^*.

Proof. It suffices to show that, given τ, P^*, $\underline{X} = \underline{x}$,

$$EFF(\psi^B|\underline{x}) \ge EFF(\psi'|\underline{x}), \quad \forall \psi' \in \mathcal{D}' .$$

Suppose $\psi^B(x) = (0,\ldots,\underbrace{\nu,1,\ldots,1}_{s \text{ terms}})$ $\quad 0 \le \nu < 1,\ 1 \le s \le k-1.$

By theorem 2.3.1 there exists $c \ge 0$ such that

$$\sum_{i=1}^{k} \psi^B_{(i)}(x) + c = \sum_{i=1}^{k} \psi'_{(i)}(x) .$$

If $0 \leq c < 1$, then

$$\mathrm{EFF}(\psi'|\underset{\sim}{x}) = \frac{\sum_{i=1}^{k} \psi'_{(i)}(x)p_{[i]}(x)}{\sum_{i=1}^{k} \psi'_{(i)}(\underset{\sim}{x})}$$

$$\leq \frac{\sum_{i=1}^{k} \psi^{B}_{(i)}(x)p_{[i]}(x) + cp_{[k-s]}(x)}{\sum_{i=1}^{k} \psi^{B}_{(i)}(x) + c}$$

$$\leq \frac{\sum_{i=1}^{k} \psi^{B}_{(i)}(\underset{\sim}{x})p_{[i]}(x)}{\sum_{i=1}^{k} \psi^{B}_{(i)}(x)}$$

$$= \mathrm{EFF}(\psi^{B}|\underset{\sim}{x}) \ .$$

If $1 \leq c = \nu' + t + (1-\nu)$, $t \geq 0$ integer, $0 < \nu' < 1$ then

$$\sum_{i=1}^{k} \psi'_{(i)}(x)p_{[i]}(\underset{\sim}{x}) = \sum_{i=k-s+1}^{k} \psi^{B}_{(i)}(x)p_{[i]}(x) + \nu' p_{[k-s-t+1]}(x)$$

$$+ p_{[k-s-t]}(x) + \ldots + (1-\nu)p_{[k-s]}(x)$$

$$\leq \sum_{i=1}^{k} \psi^{B}_{(i)}(x)p_{[i]}(x) + cp_{[k-s]}(x)$$

hence by the same argument as above we have

$$\mathrm{EFF}(\psi'|x) \leq \mathrm{EFF}(\psi^{B}|x) \ .$$

Since $\underset{\sim}{x}$ is arbitrary, the result holds for all $\underset{\sim}{x}$.

2.5. Applications to Normal Model

Suppose we have k populations $\pi_1, \ldots, \pi_k$; population π_i has distribution $N(\mu_i, \sigma_i^2)$, where σ_i's are known and μ_i's are unknown. Assume that we

have independent observations $X_{i1},\ldots,X_{in_i}$, $i = 1,\ldots,k$. Let $\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$ and let $\underline{X} = (\bar{X}_1,\ldots,\bar{X}_k)$.

Suppose we are interested in selecting a subset containing the best (the population having the largest mean) under the posterior-P^* condition, wrt some prior $\tau = \tau(\underline{\mu})$. Then to find a Bayes-$P^*$ selection procedure is equivalent, in some sense, to finding $p_i(\underline{x})$, which is the posterior probability of the event $\{\pi_i$ is the best$\}$, given observations $\underline{X} = \underline{x}$, wrt a given prior τ, for all $i = 1,\ldots,k$.

Case I. *Assume that we have a common sample size n and a common known* variance σ^2.

Ia. Suppose we have no prior information about the unknown parameters, and use the "non-informative" (Box and Tiao (1973)) or "locally uniform" prior $p(\mu_i) \propto c$ for each population.

The posterior density function g_i of μ_i, given x is the normal density with mean x_i and variance σ^2/n, i.e.,

$$g_i(\mu_i|x) = \frac{\sqrt{n}}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{n(\mu_i - x_i)^2}{2\sigma^2}\right).$$

Hence

$$p_{[i]}(x) = P(\mu_{(i)} = \mu_{[k]} | X = x)$$

$$= \int_{-\infty}^{\infty} \prod_{j\neq i} \Phi\left(t + \frac{\sqrt{n}}{\sigma}(x_{[i]} - x_{[j]})\right) d\Phi(t)$$

$$i = 1,\ldots,k.$$

Here $\mu_{(i)}$ is the quantity corresponding to the i^{th} largest observation $x_{[i]}$.

Ib. If μ_i's are independent and have the identical prior distribution $N(\theta_0,\sigma_0^2)$ and $X_i|\mu_i \sim N(\mu_i,\sigma_1^2/n)$, then it is well known that the posterior density funciton g_i of μ_i, given $\underline{X} = \underline{x}$ is

$$g_i(\mu_i|\underline{x}) \sim N(\bar{\theta}_{x_i},\xi^2) \text{ with SIP property}$$

where

$$\bar{\theta}_{x_i} = \xi^2(\sigma_0^{-2}\theta_0 + n\sigma_1^{-2}x_i)$$

$$\xi^2 = (\sigma_0^{-2} + n_1\sigma_1^{-2})^{-1}.$$

Hence

$$p_{[i]}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j\neq i} \Phi(t + \xi n\sigma_1^{-2}(x_{[i]} - x_{[j]})d\Phi(t).$$

The last expression for $p_{[i]}(\underline{x})$ is the same as that for the non-informative prior whenever $\sigma_0 \to \infty$.

Since $p_{[i]}(\underline{x}) = p_{[i]}(\underline{x} + \underline{b})$ and since the normal distribution has the strictly SIP, it follows that ψ^B and ψ^B_{NR} are "just" a.e. and translation-invariant in both case Ia and Ib.

<u>Case II</u>. Variance σ_i's are known but σ_i's and n_i's are not all equal.

IIa. Using the non-informative prior $p(\mu_i) \propto c$, $i=1,\ldots,k$, we have

$$p_{(i)}(\underline{x}) = \int_{-\infty}^{\infty} \prod_{j\neq i} \Phi(t \frac{v_{(i)}}{v_{(j)}} + \frac{x_{[i]}-x_{[j]}}{v_{(j)}})d\Phi(t)$$

where $v_{(i)} = \frac{\sigma_{(i)}}{n_{(i)}}$ $i = 1,\ldots,k$. $p_{(i)}$, $\sigma_{(i)}$ and $n_{(i)}$ are corresponding to $x_{[i]}$ and we have the following theorem.

Theorem 2.5.1. $p_{(i)}(\underline{x})$ is non-decreasing in i, i.e., $p_{(i)}(\underline{x}) = p_{[i]}(x)$.

Remark 2.5.1. From the above formula of $p_{(i)}(\underline{x})$, it is easy to see, increasing the sample size of the non-best populations will increase the probability that the best population to be selected, however, before doing this, we don't know which one is the best one.

In this case ψ^B and ψ^B_{NR} are "just" a.e. and translation-invariant.

Case III. Assume that priors are independent but not identical normal distributions, namely, $\mu_i \sim N(\theta_i, \sigma^2_{0i})$, where θ_i's are not all equal; if the conditional distribution of X_i, given μ_i, is $N(\mu_i, \frac{\sigma^2_{1i}}{n_i})$, then the posterior density of μ_i, given $X_i = x_i$ is $g_i(\mu_i | x_i)$, which is the probability density function of normal distritubion $N(\bar{\theta}_{x_i}, \xi^2_i)$ where

$$\bar{\theta}_{x_i} = \xi^2_i(\sigma^{-2}_{0i}\theta_i + n_i\sigma^{-2}_{1i}x_i)$$

$$\xi^2_i = (\sigma^{-2}_{0i} + \sigma^{-2}_{1i}n_i)^{-1} .$$

Hence we have

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi[t \frac{\xi_i}{\xi_j} + \frac{1}{\xi_j}(\bar{\theta}_{x_i} - \bar{\theta}_{x_j})]d\Phi(t).$$

If $\sigma_{0i} = \sigma_0$, $\sigma_{1i} = \sigma_1$ and $n_i = n$, $i = 1,\ldots,k$, then

$$\xi_i = \xi = (\sigma_0^{-2} + \sigma_1^{-2}n)^{-1} \qquad i = 1,\ldots,k$$

and

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi[t + \xi(\frac{\theta_i - \theta_j}{\sigma_0^2} + \frac{n(x_i - x_j)}{\sigma_1^2})]d\Phi(t) .$$

Case IV. The General Normal Model

Here we consider a more general prior. Suppose we have k populations, common sample size n for each population, and common known variance $\sigma^2 > 0$. The observation can reduce to $\underline{X} = (X_1, \ldots, X_k)$ where $X_i = \sum_{i=1}^{n} X_{ij}/n$, by sufficiency.

The "Normal Model" is defined as follows:

(a) $$\underline{X} \mid \underline{\mu} \sim N(\underline{\mu}, qI), \quad q = \frac{\sigma^2}{n}$$

where I is the $k \times k$ identity matrix.

So the X's are (conditionally) independent when $\underline{\mu}$ is given.

(b) $$\underline{\mu} \sim N(\theta_0 \underline{1}, \gamma I + tU)$$

where $\theta_0 \in \mathbb{R}$, $\gamma > 0$, $t > -\frac{\gamma}{k}$,

$$\underline{1} = (1, \ldots, 1) \quad \text{and} \quad U = \underline{1}'\underline{1}\ .$$

Here $\gamma > 0$ and $t > -\frac{\gamma}{k}$ are necessary and sufficient for $\gamma I + tU$ to be positive definite. This model was chosen by Chernoff and Yahav (1977) ($t > 0$), Gupta and Hsu (1978) and Miescke (1979).

By (a) and (b) we get the posterior distribution of $\underline{\mu}$, given $\underline{X} = \underline{x}$, and the distribution of $\underline{X}$ as follows:

$$\underline{\mu} \mid \underline{x} \sim N(\underline{\theta}, aI + bU)$$

where

$$\underline{\theta} = \gamma(q+\gamma)^{-1}\underline{x} + qt((q+\gamma)(q+\gamma+kt))^{-1}\,\underline{x}\,U + q(q+\gamma+kt)^{-1}\theta_0\,\underline{1}$$

$$a = \gamma q(q + r)^{-1}$$

$$b = q^2 t(q+\gamma)^{-1}(q+\gamma+kt)^{-1}$$

$$\underline{X} \sim N(m\underline{1},(q + \gamma)I + tU)$$

Lemma 2.5.1. Let $\underline{Y} \sim N(\underline{\mu} + \rho\underline{1}, aI + bU)$ with $\underline{\mu} \in \mathbb{R}^k$, $\rho \in \mathbb{R}$, $a > 0$ and $b > -a/k$. Then there exists a random vector $\underline{Z} \sim N(\underline{\mu}, aI)$ such that $h(Y) = h(Z)$ everywhere for every translation-invariant $h: \mathbb{R}^k \to \mathbb{R}^k$.

Proof. (See Miescke (1979)).

With this lemma, it is easy to get

$$p_i(\underline{x}) = P(\mu_i = \mu_{[k]}|\underline{x})$$

$$= \int I_{\{\mu_i=\mu_{[k]}\}} d\Phi_{((\frac{\gamma}{q+\gamma})\underline{x},\ \frac{\gamma q}{q+\gamma} I)}(\underline{\mu})$$

where $\Phi_{(\underline{\mu},V)}$ is the normal distribution with mean $\underline{\mu}$ and variance-covariance matrix V.

We can rewrite $p_i(x)$ as

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j\neq i} \Phi(t + (\frac{\gamma}{q(q+\gamma)})^{\frac{1}{2}}(x_i - x_j))d\Phi(t) .$$

Let $\gamma = \sigma_0^2$, $q = \sigma^2/n$, we have

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j\neq i} \Phi(t + (\frac{\sigma_0^2}{\frac{\sigma^2}{n}(\frac{\sigma^2}{n} + \sigma_0^2)})^{\frac{1}{2}}(x_i - x_j))d\Phi(t).$$

The above expression is exactly the same as that of the independent prior Case I, Ib.

Case V. Under normal assumption as before, but suppose σ_i's are unknown and that neither σ_i's nor n_i's are all equal.

Suppose we have no prior information about (μ,σ), for each individual population π_i assign prior $p(\mu_i,\sigma_i)\propto\sigma_i^{-1}$ then we have (See Box and Tiao (1973)) that the posterior density of μ_i, given $X_i = x_i = (x_{i1},\ldots,x_{in_i})$ is

$$p(\mu_i|x_i) = \frac{(s_i/\sqrt{n_i})^{-1}}{B(\frac{1}{2}\nu_i,\frac{1}{2})\sqrt{\nu_i}}\left[1+\frac{n_i(\mu_i-x_i)^2}{\nu_i s_i^2}\right]^{-\frac{1}{2}(\nu_i+1)}$$

where s_i^2 is the sample variance, $B(\cdot,\cdot)$ is a Beta function and $\nu_i = n_i - 1$. Hence

$$p(t_i = \frac{\mu_i - x_i}{s/\sqrt{n_i}} | x_i) = \frac{1}{B(\frac{1}{2}\nu_i,\frac{1}{2})\sqrt{\nu_i}}\left(1+\frac{t_i^2}{\nu_i}\right)^{-\frac{1}{2}(\nu_i+1)},$$

which is the density of the student's t distribution with $\nu_i (= n_i - 1)$ degrees of freedom.

Using this result we can write the formula of $p_i(x)$ by

$$p_i(\underset{\sim}{x}) = P(\mu_i > \mu_j, \quad \forall j \neq i | x)$$

$$= \int \prod_{j\neq i} T\nu_j\left(t\,\frac{s_i/\sqrt{n_i}}{s_j/\sqrt{n_j}} + \frac{x_i - x_j}{s_j/\sqrt{n_j}}\right) dT\nu_i(t)$$

where $\quad \nu_i = n_i - 1, \quad i = 1,\ldots,k$

$$\nu_i s_i = \sum_{r=1}^{n_i} (x_{ir} - x_i)^2$$

$T\nu_i$ is the c.d.f. of t distribution with ν_i degrees of freedom. When ν_i's are large, t distribution approaches normal distribution, hence, for large n_i, $i = 1,\ldots,k$, we can replace T by Φ.

Case VI. Suppose we are interested in finding a subset which contains the population with the smallest variance; i.e., we define the best population as the one with the smallest variance, and suppose that we have no prior information about σ. In this case, it is reasonable to assume that

$$p(\mu,\sigma) \propto \sigma^{-1}, \text{ if } \mu \text{ is unknown}$$

$$p(\sigma) \propto \sigma^{-1}, \text{ if } \mu \text{ is known.}$$

Let

$$\nu_j = n_j,\ \nu_j S_j^2 = \sum_{r=1}^{k} (X_{jr}-\mu)^2 \text{ if } \mu \text{ is known}$$

$$\nu_j = n_j - 1,\ \nu_j S_j^2 = \sum_{r=1}^{k} (X_{jr}-\bar{X}_j)^2 \text{ if } \mu \text{ is unknown, } n_j > 1 \quad j = 1,2,\ldots,k$$

$$S^2 = (S_1^2,\ldots,S_k^2),\ X = (X_{11},\ldots,X_{1n_1},\ldots,X_{kn_k})$$

and Y_ν be the random variable with c.d.f. χ_ν^2 which is the χ^2 distribution with ν degrees of freedom.

Then for either case (μ known or unknown), we have

$$p_i(x) = P(\sigma_i^2 = \sigma_{[1]}^2 \mid X = x)$$

$$= P(\sigma_i^2 \le \sigma_j^2,\ \forall\, j \ne i \mid X = x)$$

$$= P\left(\frac{\nu_j s_j^2}{\sigma_j^2} \le \frac{\nu_i s_i^2}{\sigma_i^2}\left(\frac{\nu_j s_j^2}{\nu_i s_i^2}\right),\ \forall\, j \ne i \mid S^2 = s^2\right)$$

$$= P\left(Y_{\nu_j} \le Y_{\nu_i}\left(\frac{\nu_j s_j^2}{\nu_i s_i^2}\right),\ \forall\, j \ne i \mid S^2 = s^2\right)$$

$$= \int_0^\infty \prod_{j \ne i} \chi^2_{\nu_j}\left(u\,\frac{\nu_j s_j^2}{\nu_i s_i^2}\right) d\chi^2_{\nu_i}(u)$$

$$= \int_0^\infty \prod_{j \neq i} \Phi_\nu^2\left(u \frac{s_j^2}{s_i^2}\right) d\Phi_\nu^2(u) \quad \text{if } n_1 = \dots = n_k = n.$$

With these $p_1(x), \dots, p_k(x)$ we can apply Bayes-P^* rules ψ^B and ψ^B_{NR} easily.

Lemma 2.5.2. In Case VI, ψ^B and ψ^B_{NR} are just * a.e. and (scale) translation invariant.

* Here the definition of the "just" property for a selection rule is

$$\psi_i(s^2) \geq \psi_i(s^{2'}) \text{ if } s_i^2 \geq s_i^{2'},\ s_j^2 \leq s_j^{2'},\ \forall j \neq i .$$

2.6. Comparison of Selection Rules ψ^B and ψ^M in the Normal Location Parameter Case

We have k normal populations with a common known variance σ^2 and common sample size n. For this case Gupta (1956) proposed and studied the procedure ψ^M.

ψ^M: Select π_i iff $X_i \geq X_{[k]} - d\frac{\sigma}{\sqrt{n}}$, $i = 1, \dots, k$ where $d = d(k, P^*)$ is to be determined by

$$\inf_{\theta \in \Omega} P(CS|\psi^M) = P^*$$

and Ω is the parameter space.

We will show that $\psi^M \in \mathcal{A}_{NR}(\cdot, P^*)$ where τ is the locally uniform prior distribution. For fixed P^* and k, let d be determined by

$$\int_{-\infty}^{\infty} \Phi^{k-1}(t + d)d\Phi(t) = P^* . \tag{2.6.1}$$

Let

$$\mathcal{X} = \{\text{all possible observed values}\} = \mathbb{R}^k$$

$$\mathcal{X}_1 = \{x \in \mathcal{X} \mid x_{[k]} - d\frac{\Delta}{\sqrt{n}} \le x_{[1]}\}$$

$$\mathcal{X}_i = \{x \in \mathcal{X} \mid x_{[i-1]} < x_{[k]} - d\frac{\Delta}{\sqrt{n}} \le x_{[i]}\}, \quad 2 \le i \le k$$

$$\mathcal{X}_i^{(1)} = \{x \in \mathcal{X} \mid x_{[1]} = x_{[i-1]} < x_{[k]} - d\frac{\Delta}{\sqrt{n}} \le x_{[i]}\} \subset \mathcal{X}_i$$

$$\mathcal{X}_i^{(2)} = \{x \in \mathcal{X} \mid x_{[1]} = x_{[i-1]} < x_{[k]} - d\frac{\Delta}{\sqrt{n}} = x_{[i]} = x_{[k-1]}\} \subset \mathcal{X}_i^{(1)}$$

then we have the following theorem.

Theorem 2.6.1. Given a number $P^*(\frac{1}{k} < P^* < 1)$ and locally uniform prior for each population π_i, $\underset{\sim}{X} = \underset{\sim}{x} \in \mathcal{X}_i$, then

$$P(CS|\delta^M, \underset{\sim}{X} = \underset{\sim}{x}) \ge q^*(i)$$

where

$$q^*(i) = \frac{k-i}{k-1}(1 - P^*) + P^* .$$

Hence

$$\delta^M \in \mathcal{D}_{NR}(\Delta, P^*) .$$

Proof. It is sufficient to show that

$$\inf_{x \in \mathcal{X}_i} \sum_{j=i}^{k} p_{[j]}(x) = q^*(i) = \frac{k-i}{k-1}(1 - P^*) + P^*.$$

Since $x \in \mathcal{X}_i$,

$$P(CS|\delta^M, x) \ge \inf_{x \in \mathcal{X}_i} P(CS|\delta^M, x)$$

$$= \inf_{x \in \mathcal{X}_i} \sum_{j=i}^{k} p_{[j]}(x) .$$

Without loss of generality we can assume $\frac{\sigma}{\sqrt{n}} = 1$.

Since

$$\sum_{i=1}^{k} p_{[i]}(\underline{x}) = 1 \quad \forall\, \underline{x} \in \mathfrak{X}, \text{ and } \forall\, \ell < i$$

$p_{[\ell]}(\underline{x})$ is nonincreasing for all $x_{[j]}$, $j \leq i - 1$, we have

$$\inf_{x \in \mathfrak{X}_i} \sum_{\ell=i}^{k} p_{[\ell]}(x) = \inf_{\underline{x} \in \mathfrak{X}_i^{(1)}} \sum_{\ell=i}^{k} p_{[\ell]}(x)$$

$$= 1 - \sup_{x \in \mathfrak{X}_i^{(1)}} \sum_{\ell=1}^{i-1} p_{[\ell]}(x)$$

$$= 1 - \sup_{x \in \mathfrak{X}_i^{(1)}} \sum_{\ell=1}^{i-1} \int_{-\infty}^{\infty} \prod_{j \neq \ell} \Phi(t + x_{[\ell]} - x_{[j]})d\Phi(t)$$

$$= 1 - \sup_{\underline{x} \in \mathfrak{X}_i^{(1)}} \sum_{\ell=1}^{i-1} \int_{-\infty}^{\infty} \{ \prod_{j \geq i} \Phi(t + x_{[\ell]} - x_{[j]})$$

$$\cdot \prod_{\substack{j \neq \ell \\ j < i}} \Phi(t + x_{[\ell]} - x_{[j]})\} d\Phi(t)$$

$$= 1 - \sup_{x \in \mathfrak{X}_i^{(1)}} \sum_{\ell=1}^{i-1} \int_{-\infty}^{\infty} \{ \prod_{j \geq i} \Phi(t + x_{[\ell]} - x_{[j]})$$

$$\cdot \Phi^{i-2}(t)\} d\Phi(t)$$

$$= 1 - \sum_{\ell=1}^{i=1} \int_{-\infty}^{\infty} \Phi(t - d)\Phi^{k-2}(t)d\Phi(t) \qquad (2.6.?)$$

$$= 1 - (i - 1) \int_{-\infty}^{\infty} \phi(t-d)\Phi^{k-2}(t)d\Phi(t)$$

$$= (k - i) \int_{-\infty}^{\infty} \Phi^{k-2}(t)\Phi(t - d)d\Phi(t)$$

$$+ \int_{-\infty}^{\infty} \Phi^{k-1}(t + d)d\Phi(t) \qquad (2.6.3)$$

The superimum of (2.6.3) occurs when $\underline{x} \in \mathcal{X}_i^{(2)}$. The last equality follows from the identity

$$(k - 1) \int \Phi^{k-2}(t)\Phi(t - d)d\Phi(t)$$

$$= 1 - \int \Phi^{k-1}(t + d)d\Phi(t),$$

which can be shown by the integration by parts. By (2.6.1), the second term of (2.6.3) equals P*; then use the integration by parts to the first term of (2.6.2), we get

$$\inf_{\underline{x} \in \mathcal{X}_i} \sum_{\ell=i}^{k} p_{[i]}(\underline{x}) = \frac{k-i}{k-1} [1 - P^*] + P^* \qquad (2.6.4)$$

$$= q^*(i) .$$

Remark 2.6.1. If the procedure ψ^M selects $\pi_{(k)}$ only, i.e. $\underline{X} = \underline{x} \in \mathcal{X}_k$, then by Theorem 2.6.1 we have $p_{[k]}(x) \geq P^*$ so that ψ^B or ψ^B_{NR} selects $\pi_{(k)}$ only. But the converse is not necessarily true.

Remark 2.6.2. For the case k = 2, $\psi^B_{NR} = \psi^M$ a.e. For any given $\underline{X} = \underline{x}$; if $x \in \mathcal{X}_2$, then $p_{[2]}(x) > P^*$, hence ψ^M and ψ^B_{NR} select the population $\pi_{(2)}$ associated $x_{[2]}$. If $x \in \mathcal{X}_1$, and $x_{[2]} - d \frac{\sigma}{\sqrt{n}} < x_{[1]}$ then ψ^M and

ψ_{NR}^{B} select both populations π_1 and π_2. Since

$$P(X_{[2]} - d\frac{\sigma}{\sqrt{n}} = X_{[1]}) = 0,$$

we have $\psi_{NR}^{B} = \psi^{M}$ a.e. .

Remark 2.6.3. The above Theorem and Remark 2.6.1 gives us a lower bound on the value of $\sum_{\ell=i}^{k} P_{[\ell]}(\underline{x})$, over all $x \in \mathcal{X}_i$. The exact value of the difference of the selected sizes between ψ^{M} and ψ^{B} depends on the observations.

2.7. Applications to Select $\max_{1<i<k} \theta_i$, $\theta_i = \frac{\mu_i - a}{\sigma_i}$ for Normal Distribution $N(\mu_i, \sigma_i^2)$, $i = 1,\ldots,k$

Let $\pi_1,\ldots,\pi_k$ be k independent normal populations with mean μ_i and variance σ_i^2, both μ_i and σ_i are unknown. For the goal of finding a random subset which contains the population with maximum $\theta_i = \frac{\mu_i - a}{\sigma_i}$ for some given constant a, we assume that apriori (μ_i, σ_i), $i = 1,\ldots,k$ are independent. Suppose we have n_i independent observations $X_{i1},\ldots,X_{in_i}$ from π_i, and let $\bar{X}_i$ be their sample mean, $i = 1,\ldots,k$.

Let $Y_1,\ldots,Y_n$ be i.i.d. $\sim N(\mu, \sigma^2)$. If no prior information is available to (μ, σ), we could assign a locally uniform prior $p(\mu, \sigma) \propto \sigma^{-1}$ to (μ, σ), (see Box and Tiao (1973)). And the posterior joint distribution of $\mu' = \mu - a$ and σ, given observations $Y = y = (y_1,\ldots,y_n)$ is given by

$$P(\mu', \sigma|y) = k\sigma^{-(n+1)} \exp \{-\frac{1}{2\sigma^2} [\nu s^2 + n(y' - \mu')^2]\}$$

where

$$\left.\begin{aligned} y' &= y - a, \quad y = \sum_{i=1}^{n} y_i/n \\ \nu s^2 &= \sum (y_i - y)^2, \quad \nu = n - 1 \\ k &= \sqrt{\frac{n}{2\pi}}\, [\tfrac{1}{2}\, \Gamma(\tfrac{\nu}{2})]^{-1} (\tfrac{\nu s}{2})^{\frac{\nu}{2}}. \end{aligned}\right\} \qquad (2.7.1)$$

Let $\zeta = \sqrt{n}\, (\mu-a)/\sigma$, with (2.7.1) the posterior distribution of ζ, given $Y = y$ is

$$p(\zeta|Y = y) = p(\zeta|t)$$

$$= \{2^{\frac{\nu}{2}-1}\, \Gamma(\tfrac{\nu}{2})\}^{-1} (\frac{\nu}{\nu+t^2})^{\frac{\nu}{2}} \exp \{-\frac{1}{2} (\frac{\zeta^2\nu}{\nu+t^2})\}\, \Gamma(\nu) I_{\nu-1}(\frac{t\zeta}{(\nu+t^2)^{1/2}})$$

where

$$t = \sqrt{n}(y - a)/s, \quad \nu = n - 1$$

$$I_\nu(x) = \int_0^\infty (\sqrt{2\pi}\, \Gamma(\nu))^{-1} u^\nu \exp \{-\frac{1}{2} (u + x)^2\} du.$$

Now, let $p(\mu_i, \sigma_i) \propto \sigma_i^{-1}$ be the assigned locally uniform prior to (μ_i, σ_i). Then let $x = (x_{11}, \ldots, x_{1n_1}, \ldots, x_{kn_k})$, we have

$$p_i(x) = P(\theta_i = \theta_{[k]} | x)$$

$$= P(\frac{\mu_i - a}{\sigma_i} = \max_{1 \le j \le k} (\frac{\mu_j - a}{\sigma_j}) | x)$$

$$= P\left(\sqrt{\frac{n_j}{n_i}}\,\xi_i > \xi_j \quad \forall\, j \neq i \mid \underline{t}\right)$$

$$= \int \prod_{j \neq i} G_{\xi_j}\left(\sqrt{\frac{n_j}{n_i}}\, z \mid \underline{t}\right) d\, G_{\xi_i}(z \mid \underline{t}) \qquad (2.7.2)$$

$$= \int \prod_{j \neq i} G_{\xi_j}(z \mid \underline{t})\, d\, G_{\xi_i}(z \mid \underline{t}) \quad \text{if } n_1 = \ldots = n_k = n,$$

where G_ξ is the posterior c.d.f. of ξ, given x or $\underline{t}$.

By (2.7.2), the Bayes-P* procedure is completely specified.

If the prior distribution for (μ, σ) is the conjugate distribution (see Raiffa and Schlaifer (1960)), then

$$p(\mu, \sigma) \propto \exp\left\{-\frac{1}{2\sigma^2} n'(\mu - m')^2\right\} \frac{1}{\sigma} \cdot \exp\left\{-\frac{1}{2\sigma^2}(\nu' v')\right\}$$
$$= p(\mu \mid \sigma)\, p(\sigma)$$

that is

$$p(\mu \mid \sigma) \sim N(m', \sigma^2/n'), \quad n' > 0$$

$$p(\sigma) \sim \frac{\nu' v'}{\sigma^2} \sim \chi^2_{\nu'}, \quad v', \nu' > 0.$$

Let

$$x' = \frac{nx + n'm'}{n + n'}, \quad x \text{ is the sample mean.}$$

$$u^2 = \{(n - 1)s^2 + \nu' v' + [(nn')/(n + n')](x - m')^2\}/\nu^*$$

$$\nu^* = (n - 1) + \nu' + 1$$

$$\xi^* = (n + n')^{1/2}(\mu - a)/\sigma$$

$$t^* = (n + n')^{1/2}(x' - a)/u,$$

the posterior distribution of ξ^*, given x is $p(\xi^* \mid x) = p(\xi^* \mid t^*)$ which has the same form as $p(\xi \mid t)$, but replace ξ, t, ν by ξ^*, t^*, ν^*.

Thus for the conjugate prior case, we get

$$p_i(\cdot) \quad p(\; _i \; _{[k]} \; \cdot)$$

$$\int \; _{j \neq i} \; G_{\, \overset{\star}{j}} \; \Big/ \; _i \; z \; t^{\star}) d \, G_{\, \overset{\star}{i}}(z \; t^{\star})$$

$$\int \; _{j \neq i} \; G_{\, \overset{\star}{i}} \; (z \; t^{\star}) d \; G_{\, \overset{\star}{i}}(z \; t^{\star}) \quad \text{if } n_i$$

where $G_{\star}$ is the posterior c.d.f. of * given x or t.

Note that (2.7.3) has the same form as (2.7.2), but replace , t by *, $t^{\star}$.

2.8. Applications to Poisson Distributions and Poisson Processes

2.8.1. Poisson Distributions Case

Suppose that $_1, \ldots, _k$ are k independent Poisson populations, where the independent observations $X_{i1}, \ldots, X_{in_i}$ from $_i$ have the Poisson density with parameter $_i$; denoted by $P(\cdot \; _i)$, i = 1,...,k.

Let $Y_1, \ldots, Y_n$ be i.i.d. with $p(\cdot \;)$. If we use non-informative prior $p(\cdot) \; ^{-1/2}$ (Box and Tiao (1973)), then given Y = y = ($y_1, \ldots,$ we have the posterior density as follows:

$$p(\cdot \mid y) = c \; ^{ny -} \; \exp(-n \cdot)$$

where

$$y = \frac{1}{n} \sum_{i=1}^{n} y_i \text{ and } c = n^{ny + \frac{1}{2}} [\; (ny + \tfrac{1}{2})]^{-1}.$$

We see that $2n\lambda|y \sim \chi^2_{2ny+1}$, the chi-square distribution with $2ny+1$ degrees of freedom. Hence by using non-informative prior $p(\lambda_i) \propto \lambda_i^{-1/2}$ for each population π_i, we have

$$p_i(x) = p(\lambda_i = \lambda_{[k]}|x)$$

$$= \int_{-\infty}^{\infty} \prod_{j\neq i} \chi^2_{\nu_j}\left(z\,\frac{n_i}{n_j}\right) d\chi^2_{\nu_i}(z)$$

where

$$\nu_i = 2n_i\bar{x}_i + 1, \quad \bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i .$$

If $n_1 = \cdots = n_k$, then

$$p_i(x) = \int_0^{\infty} \prod_{j\neq i} \chi^2_{\nu_j}(z)\, d\chi^2_{\nu_i}(z).$$

With $p_i(x)$, $i = 1,\ldots,k$, we can apply Bayes-P* selection rules ψ^B and ψ^B_{NR} easily to select a subset which contains the population with the largest parameter λ. On the other hand, if we are interested in selecting the population with the smallest parameter λ, then

$$p_i(x) = \int_0^{\infty} \prod_{j\neq i} \left[1 - \chi^2_{\nu_i}\left(z\,\frac{n_j}{n_i}\right)\right] d\chi^2_{\nu_j}(z)$$

$$= \int_0^{\infty} \prod_{j\neq i} \left[1 - \chi^2_{\nu_j}(z)\right] d\chi^2_{\nu_i}(z) \quad \text{if } n_1 = \cdots = n_k .$$

In this case, the simulation results for selection procedures ψ^B and ψ^B_{NR} are tabulated on Table VII.

2.8.2. Poisson Processes Case

Suppose we have k independent Poisson processes $\{X^{(1)}(t)\},\ldots,\{X^{(k)}(t)\}$ with expected arrival times equal to $\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_k}$, respectively. Hence for the processes $\{X^{(i)}(t)\}$, the probability that there are m_i arrivals until time t_i is

$$p(X^{(i)}(t_i) = m_i | \lambda_i, t_i) = \frac{(t_i\lambda_i)^{m_i}}{m_i!} e^{-t_i\lambda_i}$$

$$i = 1, 2, \ldots, k, \quad m_i = 0, 1, 2, \ldots$$

If there exists no prior information, then we use the non-informative prior $p(\lambda_i) \propto \lambda_i^{-1/2}$ for all processes. Therefore, we get the posterior density function of λ_i, given (m_i, t_i) as follows:

$$p(\lambda_i | X^{(i)}(t_i) = m_i, t_i) = p(\lambda_i | m_i, t_i)$$

$$= \frac{(t_i\lambda_i)^{m_i + \frac{1}{2} - 1}}{\Gamma(m_i + \frac{1}{2})} t_i e^{-t_i\lambda_i}.$$

Thus $2t_i\lambda_i$ has χ^2 distribution with $2m_i+1$ degrees of freedom, given the number m_i of arrivals before time t_i.

Let $m = (m_1,\ldots,m_k)$ and $t = (t_1,\ldots,t_k)$, then it can be shown that the Poisson process $\{X^{(i)}(t)\}$ has the maximum parameter (or minimum mean waiting time), given (m, t) is

$$p_i(\underline{m}, \underline{t}) = \int_0^\infty \prod_{j \neq i} \chi^2_{2m_j+1}\left(y \frac{t_j}{t_i}\right) d\, \chi^2_{2m_i+1}(y) \quad i = 1,\ldots,k. \qquad (7.8.1)$$

Here we list two special cases which are of interest.

(a) Observations of all processes are obtained in a common time interval $[s_i, t + s_i]$. Since Poisson process is stationary, we can assume that $s_i = 0$, and $t_1 = \ldots = t_k = t$. In this case

$$p_i(\underline{m}, \underline{t}) = \int_0^\infty \prod_{j \neq i} \chi^2_{2m_j+1}(y) d\, \chi^2_{2m_i+1}(y)$$

which is independent of t.

(b) All m_i's are equal, i.e. we fix m first, then get observations t. Hence

$$p_i(m, \underline{t}) = \int_0^\infty \prod_{j \neq i} \chi^2_{2m+1}\left(y \frac{t_j}{t_i}\right) d\, \chi^2_{2m+1}(y).$$

There is an alternative way to approach the cases (a) and (b). Let T_i be the waiting time of the nth arrival in the ith process, then T_i has a gamma distribution with density given by

$$p(t) = \frac{\lambda_i}{\Gamma(m_i)} (\lambda_i t)^{m_i - 1} e^{-\lambda_i t} \quad t > 0.$$

If we have only non-informative prior $p(\lambda) \propto \lambda^{-1/2}$; then, given m_i and t_i, $2 t_i \lambda_i$ has posterior distribution $\chi^2_{2m_i+1}$, therefore the formula of $p_i(\underline{m}, \underline{t})$ we get here is exactly the same as before.

Remark 2.8.1. Under non-informative prior, in comparing the subset selection problem in k Poisson distributions with the problem in k Poisson processes, it is easily seen that Poisson distributions model is a special case of Poisson processes model, namely, $t_i = n_i$ where n_i denotes the sample size of the ith Poisson population.

2.8.3. Relation Between Selection from Poisson Processes and Selection from Populations with Gamma or Exponential Distribution

Suppose we have k independent populations, the ith population having the gamma distribution with parameters $\alpha = m_i$ (known), $\beta = 1/\lambda_i$ (unknown). Since the random variable T_i, the waiting time until m_i arrivals in a Poisson process with parameter λ_i, has a gamma distribution with parameters $\alpha = m_i$, $\beta = 1/\lambda_i$. If the m_i's are given and if the goals for both selection problems are the same, namely, to select a subset containing the population (process) with the largest parameter λ, then it is easily seen that these are identical problems. Note that in the selection problem of Poisson processes, m_i's might not be the preassigned values but are given random observations whenever t_i's are preassigned values. In this case, the selection problem of Poisson processes is different from that of the gamma distributions.

If the process associated with the minimum parameter λ (or the maximum waiting time) is the best, then the posterior probability of process $\{X^{(i)}(t)\}$ to be the best is analogous to the one obtained

before with the modifications that the integrand function

$$\prod_{j\neq i} \Phi^2_{2m_j+1}\left(y\frac{t_j}{t_i}\right)$$

of (2.8.1) is replaced by

$$\prod_{j\neq i} \left[1 - \Phi^2_{2m_j+1}\left(y\frac{t_j}{t_i}\right)\right].$$

2.9. Comparison of the Performance of ψ^B, ψ^B_{NR}, ψ^M and ψ^{MED}

Let π_i, $i = 1,\ldots,k$ be k independent populations, where π_i has the associated c.d.f. $F(x, \theta_i) = F(x - \theta_i)$ with unknown location parameter θ_i. Let $f(x, \theta_i) = f(x - \theta_i)$ be the p.d.f. The goal is to find a small (nontrival) subset which contains the best.

The following subset selection procedure ψ^{MED} based on sample medians is due to Gupta and Singh (1980).

ψ^{MED}: Select π_i if and only if $\check{X}_i \geq \check{X}_{[k]} - d$

where $\check{X}_i$ is the median of the $2m+1$ random observations from population π_i and $\check{X}_{[k]} = \max_i \check{X}_i$. The value d is determined by the following equation so that the P*-condition is met.

$$\int_{-\infty}^{\infty} G(u + d)^{k-1} g(u)du = P^*$$

where

$$g(u) = \frac{(2m+1)!}{(m!)^2} [F(u)]^m [1 - F(u)]^m f(u)$$

$$G(u) = I_{F(u)}(m + 1, m + 1)$$

$I_y(p, q)$ is the incomplete beta function.

In this section we use Monte Carlo simulation techniques to compare the performance of selection procedures ψ^B, ψ^B_{NR}, ψ^M and ψ^{MED} in the normal means problem. Because both rules ψ^M and ψ^{MED} are not based on any prior information about the unknown parameters, we assume that the prior distribution τ for both ψ^B and ψ^B_{NR} is locally uniformly distributed. Since the selection procedure ψ^M satisfies both the P*-condition and the posterior-P* condition wrt the locally uniform priors, it makes sense to compare the Bayes-P* procedures ψ^B and ψ^B_{NR} with ψ^M and compare ψ^M with ψ^{MED} in terms of efficiency which is the ratio of the probability of a correct selection to the expected selected size. For studying the robustness of these four rules, ψ^B, ψ^B_{NR}, ψ^M and ψ^{MED}, we change the true distribution to non-normal distributions, namely, the logistic, Laplace (the double exponential) and the gross error model (the contaminated distribution), but keep the selection procedure unchanged (i.e. still based on the normal assumption). The Monte Carlo simulation results for both equal distances of the parameters and slippage cases are tabulated. In the simulation study all generated random variables are adjusted to have variance 1. Each time we generate five random variables with the given distribution of each population, then apply the selection procedures. The simulation process is repeated 100 times for each random variable. The relative frequency of selecting the population π_i is

used as an approximation to the probability of selecting the population π_i. The sum of relative frequency of selecting each population π_i, $i = 1,\ldots,k$ is treated as an approximation of the expected selected size. The efficiency EFF of each selection procedure is approximated by the ratio of relative frequency of selecting the best one to the expected size. The simulation results indicate that in all cases we have the performance

$$\psi^B > \psi^B_{NR} > \psi^M .$$

It should be noted that in the above comparison of the performance, we restrict attention to these rules which satisfy the posterior-P^* condition. For small sample size, the efficiency of rule ψ^M tends to be larger than ψ^{MED} under P^*-condition.

Remark 2.9.1. The Laplace distribution has the density function

$$f(x - \theta) = \frac{1}{2} e^{-|x-\theta|}, \quad -\infty < x < \infty$$

for which the variance is 2.

The logistic distribution has the density function

$$f(x - \theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}$$

for which the variance $Var(X) = \frac{\pi^2}{3}$.

The gross error model we used has the density function

$$f(x - \theta) = (1 - \varepsilon)\phi(x - \theta) + \frac{\varepsilon}{4}\phi\left(\frac{x-\theta}{4}\right), \quad \varepsilon = .15$$

for which ϕ is the p.d.f. of $N(0,1)$ and the variance $Var(X) = (1 - \varepsilon) + \varepsilon \cdot 4^2 = 3.25$.

The efficiency of a selection procedure δ is defined by

$$EFF_{\mu}(\delta) = \frac{P_{\mu}(CS|\delta)}{E_{\mu}(S|\delta)}$$

where $E_{\mu}(S|\delta)$ is the expected selected size.

Discussion and Conclusion

For Table VIII.1 and Table VIII.2 (equal distances case) the P* is .99 and .90 respectively, the common sample size n = 5, k = [illegible]. If the k populations have normal distributions with the unknown parameter configuration ($\theta, \ldots, \theta + (k-1)\delta$), common variance 1. From both tables the performance based on either the efficiency or the expected selected size is

$$\delta^{B} \succ \delta^{B}_{NR} \succ \delta^{M}$$

if the posterior-P* condition is considered, and

$$\delta^{M} \succ \delta^{MED}$$

under the P*-condition.

When the true distributions are not normal, but the logistic, the Laplace or the gross error model, the results are very close to the normal case, hence the [illegible] rules are robust. From Table VIII.2 all [illegible] are larger than the corresponding ones in Table VIII.1. This [illegible] because the value of P* is smaller in the [illegible] table.

For Tables [illegible] and [illegible] ([illegible] case) the P* is [illegible] and .90 respectively, the common sample size [illegible], k [illegible]. If the k populations [illegible]

have normal distributions with unknown parameter configurations $(\theta,\ldots,\theta+\Delta)$, common variance 1. From both tables the performance based on either the efficiency or the expected selected size is

$$\psi^B > \psi^B_{NR} > \psi^M$$

if the posterior-P* condition is considered, and if $\Delta\sqrt{n} \geq 1$

$$\psi^M > \psi^{MED}$$

under the P*-condition.

Note that in both equal distances and slippage cases when $\Delta\sqrt{n} \geq 1$, that is the population means are not very close, the procedures ψ^B and ψ^B_{NR}, wrt the locally uniform priors, always satisfy not only the posterior-P* condition but also $P_{\underline{\theta}}(CS|\psi^B \text{ or } \psi^B_{NR}) \geq P^*$, and the expected selected size of the selection procedure ψ^B or ψ^B_{NR} is much less than the selection procedures ψ^M and ψ^{MED}. For example, in the normal equal distances case, $P^* = .99$, $k = 5$, $\Delta\sqrt{n} = 4$,

$$E(S|\psi^{MED}) - E(S|\psi^B_{NR}) = 0.382;$$

in the normal slippage case, $P^* = .99$, $k = 5$, $\Delta\sqrt{n} = 4$

$$E(S|\psi^{MED}) - E(S|\psi^B_{NR}) = 1.560.$$

TABLE III

For procedures R^B and R^B_{NR} and the parameter configurations (.5,...,.5k) of k Poisson populations, this table gives the values (based on simulation) of the probability of selecting the population with parameter .5i, i=1,...,k and the expected selected size ES. The prior distribution for each population is $p(\lambda) \propto \lambda^{-\frac{1}{2}}$.

n = 10

k		P^* = 0.99 R^B	0.99 R^B_{NR}	0.95 R^B	0.95 R^B_{NR}	0.90 R^B	0.90 R^B_{NR}	0.75 R^B	0.75 R^B_{NP}
2	1	.993	1.000	.966	.990	.950	.990	.946	.990
	2	.670	.820	.450	.670	.324	.490	.110	.190
	ES	1.663	1.820	1.415	1.660	1.274	1.480	1.055	1.180
3	1	.998	1.000	.990	1.000	.967	1.000	.929	1.000
	2	.741	.850	.388	.630	.321	.510	.171	.280
	3	.205	.330	.120	.130	.065	.110	.021	.050
	ES	1.944	2.180	1.498	1.760	1.354	1.620	1.121	1.330
4	1	.993	1.000	.981	1.000	.966	.990	.911	.950
	2	.745	.820	.560	.730	.259	.400	.175	.290
	3	.369	.490	.070	.140	.067	.100	.025	.050
	4	.075	.110	.041	.060	.024	.040	0	0
	ES	2.187	2.420	1.653	1.930	1.316	1.530	1.112	1.290
5	1	.996	1.000	.981	.990	.985	1.000	.950	.960
	2	.737	.850	.503	.650	.367	.550	.128	.260
	3	.355	.470	.102	.160	.133	.210	.013	.030
	4	.067	.090	.015	.030	.013	.020	.006	.010
	5	0	[illegible]	.006	.010	0	0	0	[illegible]
	ES	2.154	2.410	1.607	1.840	1.498	1.780	1.047	1.280

TABLE VIII. 1

Efficiency (EFF) and Expected Selected Size (ES) (based on simulation) of ψ^B, ψ^B_{NR}, ψ^M and ψ^{MED} when the unknown means of the k populations are $\theta, \ldots, \theta + (k-1)\Delta$; the common variance is 1, common sample size $n = 5$ and the prior for ψ^B and ψ^B_{NR} is locally uniformly distributed.

TABLE [illegible]

$k = 5, \quad P^* = .99$

		normal		logistic		Laplace		gross error	
		EFF	ES	EFF	ES	EFF	ES	EFF	ES
$\delta\sqrt{n} = .5$	B	.254	3.809	.250	3.963	.259	3.774	.248	3.988
	B_{NR}	.238	4.110	.233	4.290	.238	4.120	.230	4.30[illegible]
	M	.208	4.810	.207	4.840	.210	4.720	.207	4.840
	MED	.208	4.810	.202	4.940	.202	4.940	.201	4.980
$\delta\sqrt{n} = 1$	B	.333	2.977	.332	3.005	.336	2.941	.329	3.032
	B_{NR}	.305	3.280	.304	3.290	.302	3.280	.304	3.29[illegible]
	M	.250	4.000	.234	4.280	.246	4.030	[illegible]	4.29[illegible]
	MED	.232	4.310	.224	4.460	.217	4.600	.214	4.6[illegible]
$\delta\sqrt{n} = 2$	B	.541	1.847	.541	1.839	.541	1.884	.559	1.77[illegible]
	B_{NR}	.48[illegible]	[illegible].050	.481	2.080	.498	2.010	.510	1.960
	M	.417	2.400	.417	[illegible].400	.515	2.410	.437	2.290
	MED	.368	2.720	.351	2.850	.36[illegible]	2.76[illegible]	.36[illegible]	2.720
$\delta\sqrt{n} = 4$	B	.825	1.212	.855	1.169	[illegible]	1.217	.865	1.156
	B_{NR}	.730	1.370	[illegible]06	1.240	.746	1.340	.80[illegible]	1.250
	M	.676	1.480	.694	1.440	.680	1.47[illegible]	.67[illegible]	1.49[illegible]
	MED	[illegible]	[illegible]	.57[illegible]	1.730	[illegible]	1.70[illegible]	.57[illegible]	1.75[illegible]

TABLE VIII. 2

Efficiency (EFF) and Expected Selected Size (ES) (based on simulation) of ψ^B, ψ^B_{NR}, ψ^M and ψ^{MED} when the unknown means of the k populations are $\mu, \ldots, \mu + (k-1)\Delta$; the common variance is 1, common sample size $n = 5$ and the prior for ψ^B and ψ^B_{NR} is locally uniformly distributed.

TABLE [illegible]

[illegible] = 5, F^* [illegible]

	Normal		Logistic		Laplace		Gross error	
	EFF	ES	EFF	ES	EFF	ES	EFF	ES
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.332	2.447	.333	2.611
[illegible]	[illegible]	[illegible]	.313	2.780	.316	2.820	.310	3.330
[illegible]	[illegible]	4.[illegible]	.252	[illegible].930	.237	4.090	.231	4.280
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.220	4.490	.215	4.660
[illegible]	[illegible]	[illegible]	[illegible]	1.817	.516	1.814	.502	[illegible]
[illegible]	[illegible]	[illegible]	.484	[illegible]	[illegible]	2.1[illegible]	[illegible]	2.1[illegible]
[illegible]	[illegible]	[illegible]	.331	[illegible]	[illegible]	3.[illegible]	.331	2.93[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	3.4[illegible]	.292	3.420	.283	3.6[illegible]
[illegible]	[illegible]	1.27[illegible]	.77[illegible]	1.2[illegible]	.756	1.291	.786	1.227
[illegible]	[illegible]	1.49[illegible]	.67[illegible]	1.4[illegible]	.664	1.490	.71[illegible]	1.390
[illegible]	[illegible]	[illegible]	.641	[illegible]	.635	1.[illegible]	.579	1.71[illegible]
[illegible]	[illegible]	[illegible]	.44[illegible]	[illegible]	[illegible]	[illegible]	.54[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.46[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE IX. 1

[illegible] Selected Size (ES) (based on simulation) of ^{B}, $_{NR}^{B}$, ^{M} and ^{MED} when the [illegible] ulations are [illegible],..., [illegible]: the common variance is 1, common sample size $n = 5$ and the [illegible] for ^{B} and $_{NR}^{B}$ is *locally uniformly* distributed.

TABLE I[illegible]. 1

[illegible] = 5, P* = .99

		normal		logistic		Laplace		gross error	
		EFF	ES	EFF	ES	EFF	ES	EFF	ES
	B	[illegible]	4.19[illegible]	.218	4.253	.226	4.203	.220	4.361
[illegible] = .5	[illegible]	[illegible]	4.52[illegible]	.212	4.580	.214	4.530	.212	4.670
	M	.201	4.97[illegible]	.202	4.950	.200	4.900	.202	4.890
	MED	[illegible]	4.930	.202	4.940	.202	4.960	.200	5.000
	B	[illegible]	4.101	.240	4.133	.240	4.062	.231	4.264
[illegible] = 1	[illegible]	[illegible]	[illegible]	.224	4.460	.223	4.430	.216	4.580
	M	.20[illegible]	[illegible]	.203	4.930	.203	4.870	[illegible]	4.900
	MED	[illegible]	4.89[illegible]	.201	4.97[illegible]	.200	4.980	.200	5.000
	B	[illegible]	[illegible]	.274	3.645	.263	3.701	.295	3.381
[illegible]	[illegible]	[illegible]	3.970	.253	3.960	.242	4.090	.269	3.720
	M	[illegible]	4.680	.211	4.750	.214	4.680	[illegible]	4.510
	MED	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	.513	1.95	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.95[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE IX. 2

[illegible] (Eff) and Expected Selected Size (ES) (based on simulation) of $[illegible]^B$, $[illegible]^B_{NR}$, $[illegible]^M$ and $[illegible]^{MED}$ when the [illegible] simulations are [illegible] the common variance is 1, common sample size [illegible] and the [illegible] $[illegible]^B$ and [illegible] is locally uniformly distributed.

TABLE [illegible]

		[illegible]		logistic		Laplace		gross error	
		EFF	ES	EFF	ES	EFF	ES	EFF	ES
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.238	2.734	.256	[illegible]
	[illegible]	[illegible]	[illegible]	.262	[illegible]	.228	3.160	.240	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.211	4.400	.212	4.490
	ME	[illegible]	[illegible]	[illegible]	[illegible]	.202	4.840	.204	4.300
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.334	2.653	.302	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.520
	ME	[illegible]	[illegible]	[illegible]	[illegible]	.207	[illegible]	.209	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.127	.500	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.385	[illegible]	.441	[illegible]
	[illegible]	[illegible]	[illegible]	.272	[illegible]	.271	[illegible]	[illegible]	3.270
	ME	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

BIBLIOGRAPHY

BIBLIOGRAPHY

Alam, K. (197[illegible]). On a multi[illegible] decision rule. Ann. Statist., [illegible] 750-755.

Andersen, E. S. (195[illegible]). On the fluctuation of sums of rando[illegible] variables. Math. Scand., [illegible], 26[illegible]-2[illegible].

Ayer, M., Brunk, H. D., Ewin[illegible], G. M., Reid, W. T. and [illegible] (1955). An empirical distribution function for sampling with incomplete information. Ann. Math. Statist., [illegible], [illegible].

Bahadur, R. R. (1950). On the problem in the theory of [illegible]. Ann. Math. Statist. 21, 36[illegible]-[illegible].

Bahadur, R. R. and Robbins, [illegible] (1950). The problem of the [illegible] mean. Ann. Math. Statist., 21, 469-48[illegible]. Correction, [illegible] 310.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and [illegible] (197[illegible]). Statistical Inference under Order Restrictions. John Wiley, New York.

Barron, A. M. and Gupta, S. [illegible] (197[illegible]). A class of non-el[illegible] sequential multiple decision procedures. Operations Research Verfahren (Ed. Henn, Künzi and Schubert). Verlag Anton [illegible], Meisenheim am Glan, Germany, pp. 11-[illegible]7.

Bechhofer, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist., [illegible], 16-[illegible].

Bechhofer, R. E. (195[illegible]). A sequential multiple decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs. Biometrics, 14, [illegible]-[illegible].

Bechhofer, R. E., Dunnett, C. W. and Sobel, M. (1954). A two-sample multiple-decision procedure for ranking means of normal populations with a common unknown variance. Biometrika, 41, [illegible].

Bechhofer, R. E., Kiefer, J. [illegible] Sobel, M. (196[illegible]). Sequential Identification and Ranking Procedures. Univ. of Chicago Press, Chicago.

Bechhofer, R. E. and Sobel, M. (195[illegible]). A single-sample multiple decision procedure for ranking variances of normal populations (prelim. report), Abstract. Ann. Math. Statist., [illegible]

Berger, R. L. (1977). Minimax, admissible and gamma-minimax [illegible] decision rules. Ph.D. Thesis (Mimeo. Ser. No. 490). [illegible] Statist., Purdue Univ., West Lafayette, Indiana.

Berger, R. L. (1979). Minimax subset selection for loss [illegible] subset size. Ann. Statist., 7, [illegible]

Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with applications to unequal variance problems. [illegible] J. Statist., 7, 21-36.

Bickel, P. J. and Yahav, J. A. (1977). On selecting a [illegible] population. Statistical Decision Theory and Related [illegible] (Eds. S. S. Gupta and D. S. Moore). Academic, New York, [illegible]

Box, G. E. P. and Tiao, G. C. (1973). Bayesian Inference [illegible] Analysis. Addison-Wesley, Massachusetts.

Brostrom, G. (1977). An improved procedure for selecting [illegible] [illegible] better than a standard. Tech. Report 1977 [illegible] Math. and Statist., Univ. of [illegible], Sweden.

Carroll, R. J. (1974). Asymptotically nonparametric sequential selection procedures. Mimeo. Ser. No. 944. Inst. [illegible] Univ. of North Carolina, Chapel Hill.

Chernoff, H. and Yahav, J. (1977). [illegible] subset selection [illegible] employing a new criterion. Statistical Decision Theory and Related Topics - II (Eds. S. S. Gupta and D. S. Moore). [illegible] New York, pp. 93-119.

Deely, J. J. (1965). Multiple decision procedures from an empirical Bayes approach. Ph.D. Thesis (Mimeo. Ser. No. [illegible]). [illegible] Statist., Purdue Univ., West Lafayette, Indiana.

Deely, J. J. and Gupta, S. S. (1968). On the properties of subset selection procedures. Sankhyā Ser. A, 30, [illegible]

Desu, M. M. and Sobel, M. (1968). A fixed subset size approach to a selection problem. Biometrika, 55, 401-410. Corrections and amendments: 63 (1976), [illegible]

Desu, M. M. and Sobel, M. (1971). Nonparametric procedures for selecting fixed-size subsets. Statistical Decision Theory and Related Topics (Eds. S. S. Gupta and J. Yackel). Academic, New York, pp. 255-273.

[illegible]nnett, [illegible] W. ([illegible]). [illegible] comparison procedure for com[illegible] several treatments with a control. J. Amer. Statist. Assoc., [illegible], 1096-1121.

[illegible]eller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. II, [illegible] Edition. J. Wiley & [illegible], New York.

[illegible]bbons, J. D., [illegible], I. and [illegible]obel, M. (1977). Selection and [illegible] Populations. Wiley, New York.

[illegible]nanadesikan, M. (1966). [illegible] selection and ranking procedures [illegible] multivariate [illegible] populations. Ph.D. Thesis. Dept. of [illegible], [illegible] Univ., West Lafayette, Indiana.

[illegible]nanadesikan, M. and Gupta, [illegible] S. (1970). Selection Procedures [illegible] multivariate normal distributions in terms of measures of [illegible]. Technometrics, 12, 103-117.

[illegible]el, P. [illegible] and [illegible], H. ([illegible]). On selecting a subset containing the best population - A Bayesian approach. Ann. Statist. [illegible], [illegible]-[illegible].

Gupta, [illegible] (1956). On a decision rule for a problem in ranking [illegible]. Ph.D. Thesis (Mimeo. Ser. No. 150). Inst. of Statist., Univ. [illegible] North Carolina, Chapel Hill.

[illegible]upta, [illegible] (1963). On a selection and ranking procedure [illegible] populations. Ann. Inst. Statist. Math., 14, 199-[illegible].

[illegible]upta, [illegible] (1965). On some multiple decision (selection [illegible]) rules. Technometrics, 7, 225-245.

[illegible]upta, [illegible] (1967). On selection and ranking procedures [illegible] Amer. Soc. Qual. Control., pp. [illegible]-155.

[illegible]upta, [illegible] and [illegible], [illegible] ([illegible]). On [illegible] procedures for selecting populations close to [illegible].

[illegible]upta, [illegible] and [illegible]su, J.C. ([illegible]). On the performance of [illegible] selection procedures. Commun. Statist. - Simul. [illegible], [illegible] 561-591.

Gupta, [illegible] S. and Huang, [illegible] ([illegible]). On subset selection [illegible] for Poisson populations and some applications [illegible] selection problems. Applied Statistics ([illegible]), [illegible] Holland, Amsterdam, pp. [illegible].

[illegible]upta, [illegible] and Huang, [illegible] ([illegible]). On some [illegible] parametric sequential subset selection procedures. [illegible] Inference and Related Topics, Vol. [illegible] (Ed. M. [illegible]), [illegible] New York, pp. [illegible].

Gupta, S. S. and Huang, D. Y. (1976). Selection procedures for the means and variances of normal populations: unequal sample size case. Sankhyā Ser. B, 38, 112-128.

Gupta, S. S. and Kim, W. C. (1980). [illegible]-minimax and minimax decision rules for comparison of treatments with a control. To appear in Recent Developments in Statistical Inference and Data Analysis. North Holland Publishing Company, 1980, 56-72.

Gupta, S. S. and McDonald, G. C. (19[illegible]). On some classes of selection procedures based on ranks. Nonparametric Techniques in Statistical Inference (Ed. M. L. Puri). Cambridge University Press, [illegible] pp. 491-51[illegible].

Gupta, S. S., and Miescke, K. J. (19[illegible]). On subset selection procedures for ranking means of three normal distributions. [illegible] Ser. No. 78-19. Dept. of Statist. Purdue Univ., W. Lafayette, Indiana. To appear in Sankhyā Ser. A.

Gupta, S. S. and Miescke, K. J. (1979). On the [illegible] configurations in certain two-stage selection procedures. [illegible] Ser. No. 79-6) Dept. of Statist., Purdue Univ., [illegible] Indiana.

Gupta, S. S. and Nagel, K. (1971). On some contributions to [illegible] decision theory. Statistical Decision Theory and Related Topics (Eds. S. S. Gupta and J. Yackel). Academic, New York, [illegible]

Gupta, S. S., Nagel, K. and Panchapakesan, S. (197[illegible]). On the [illegible] statistics from equally correlated normal random variables. Biometrika, 60, 403-413.

Gupta, S. S. and Panchapakesan, S. (197[illegible]). On a class of subset selection procedures. Ann. Math. Statist., 43, 814-8[illegible].

Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. John Wiley, New York.

Gupta, S. S. and Singh, A. K. (1980). On rules based on sample [illegible] for selection of the largest location parameter. Commun. Statist. Theor. Meth., A9(12), 127[illegible]

Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist., [illegible] 235-244.

Gupta, S. S. and Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. Biometrika, [illegible]

Gupta, S. S. and Sobel, M. (1962b). On the smallest of several correlated F-statistics. Biometrika, 49, 509-[illegible].

Gupta, [illegible] and [illegible]ffen, [illegible] On [illegible] selection and rank[illegible] procedures with applications to multivariate populations. [illegible] in Probability and Statistics (Eds. R. C. Bose et. al.). Univer[illegible] of North Carolina Press, Chapel Hill, pp. 327-[illegible].

Holm, S. (19[illegible]). A simple sequentially rejective multiple test [illegible] dure. Scand. J. Statist., 6, [illegible]-70.

Hsiao, P. (1979). Some contributions to Gamma-minimax and empirical Bayes selection procedures. Ph.D. thesis (Mimeo. Ser. No. [illegible]). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Huang, D. Y. (1975). Some contributions to fixed sample and sequential multiple decision (selection and ranking) theory. Ph.D. thesis (Mimeo. Ser. No. 363). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Huang, W. T. (1972). Some contributions to sequential selection and ranking procedures. Ph.D. thesis (Mimeo. Ser. No. 29[illegible]). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. Psychometrika, 29, 115-129.

Lehmann, E. L. (1952). Testing multiparameter hypotheses. Ann. Math. Statist., 23, 541-552.

Lehmann, E. L. (1955). Ordered families of distributions. Ann. Math. Statist., [illegible]3, 413-428.

Lehmann, E. L. (1961). Some Model I problems of selection. Ann. Math. Statist., 32, 990-1012.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. Ann. Math. Statist., 38, 1079-1091.

McDonald, G. [illegible] (1969). On some distribution-free ranking and selection procedures. Ph.D. thesis (Mimeo. Ser. No. 17[illegible]). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Miescke, K. J. (1979). Bayesian subset selection for additive and linear loss functions. Commun. Statist. - Theor. Meth., A8(12), 1205-1226.

Mosteller, F. (1948). A k-sample slippage test for an extreme population. Ann. Math. Statist., 19, 58-65.

Nagel, K. (1970). On subset selection rules with certain optimality properties. Ph.D. thesis (Mimeo. Ser. No. 222). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

[illegible] (1985). [illegible] rule for comparing [illegible] standard. Commun. Statist., [illegible], [illegible]-526.

Naik, U. D. (1977). Some subset selection problems. Commun. Statist.-Theor. Meth., A6(10), 955-966.

Naik, U. D. (1978). On selection procedures based on the [illegible] probability. Calcutta Statist. Assoc. Bull., 27, [illegible]-40.

Paulson, E. (1949). A multiple decision procedure for certain problems in the analysis of variance. Ann. Math. Statist., 20, 95-98.

Paulson, E. (1962). A sequential procedure for comparing several experimental categories with a standard or control. Ann. Math. Statist., 33, 438-443.

Paulson, E. (1963). A sequential decision procedure for choosing one of k hypothesis concerning the unknown mean of a normal distribution. Ann. Math. Statist., 34, 549-554.

Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. Ann. Math. Statist., 35, 174-180.

Paulson, E. (1967). Sequential procedures for selection of the best one of several binomial populations. Ann. Math. Statist., 38, 117-123.

Rabinowitz, P. and Weiss, G. (1959). Tables of abscissas and weights

for numerical evaluation of integrals of the form $\int_0^\infty e^{-x} x^n f(x)\,dx$.

Math. Tables and other Aids to Comput., Vol. XIII, 68, 285-294.

Rizvi, M. H. and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest α-quantile. Ann. Math. Statist., 38, 1788-1803.

Santner, T. J. (1975). A restrict subset selection approach to ranking and selection problems. Ann. Statist., 3, 334-349.

Seal, K. C. (1955). On a class of decision procedures for ranking means of normal populations. Ann. Math. Statist., 26, 387-398.

Seal, K. C. (1957). An optimum decision rule for ranking means of normal populations. Calcutta Statist. Assoc. Bull., 7, 131-150.

Sobel, M. (1967). Nonparametric procedures for selecting the t populations with the largest α-quantiles. Ann. Math. Statist., 38, 1804-1816.

Sobel, M. and Huyett, M. J. (1957). Selecting the best one of several binomial populations. Bell System Tech. J., 36, 537-576.

Studden, W. J. (1967). On selecting a subset of k populations containing the best. Ann. Math. Statist., 38, 1072-1078.

Tamhane, A. C. and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist. - Theor. Meth., A6(11), 1003-1033.

Tamhane, A. C. and Bechhofer, R. E. (1979). A two-stage minimax procedure with screening for selecting the largest mean (II): An improved PCS lower bound and associated tables. Commun. Statist. - Theor. Meth., A8(4), 337-358.

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. Contributions to Probability and Statistics (Eds. I. Olkin and others). Stanford Univ. Press, 39, 448-485.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: Mimeograph Series #80-22 — 2. GOVT ACCESSION NO.: AD-A100 948	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Some Results in the Theory of Subset Selection Procedures	5. TYPE OF REPORT & PERIOD COVERED
	6. PERFORMING ORG. REPORT NUMBER: Mimeo. Series #80-22
7. AUTHOR(s): Hwa-Ming Yang	8. CONTRACT OR GRANT NUMBER(s): ONR N00014-[illegible]
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, Department of Statistics, West Lafayette, IN 47907	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, DC	12. REPORT DATE: August 198[illegible]
	13. NUMBER OF PAGES: 13[illegible]
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

[illegible] Procedures, Subset Selection, Stagewise Procedures, [illegible] Isotonic Regression, Isotonic [illegible] Non-Informative Prior, Posterior [illegible]

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Selection and ranking (ordering) problems in statistical inference [illegible] the classical tests of homogeneity are often [illegible] situations where the experimenter is interested in [illegible] treatments or processes with the goal of selecting [illegible] populations.

Chapter I of this thesis considers the problem of selecting [illegible] population [illegible] better than a control under [illegible]

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

an ordering prior. Here, by an ordering prior we mean that there exists a known simple or partial order relationship among the unknown parameters of the treatments (excluding the control). Three new selection procedures are proposed and studied. These procedures do meet the usual requirement that the probability of a correct selection is greater than or equal to a [illegible] number P*. Two of the three procedures use the isotonic regression over the sample means of the k treatments with respect to the given [illegible]. Tables which are necessary to carry out the selection procedure with [illegible] approach for the selection of unknown means of normal populations and [illegible] populations are given. Monte Carlo comparisons on the performance of [illegible] procedures for the normal or gamma means problem were carried out in [illegible] selected cases. The results of this study seem to indicate that the [illegible] based on isotonic estimators always have superior performance [illegible] there are more than one bad population. (In comparison with [illegible])

Chapter II deals with a new 'Bayes-P*' approach about [illegible] selecting a subset which contains the 'best' of k populations. [illegible] mean the (unknown) population with the largest unknown mean. The (non-randomized) Bayes-P* rule refers to a rule with minimum [illegible] in the class of (non-randomized) rules which satisfy the condition that the posterior probability of selecting the best is at least equal to P*. Given the priors of the unknown parameters, two 'Bayes-P*' subset selection procedures δ^B and δ^B_{NR} (randomized and non-randomized, respectively) [illegible] functions are obtained and compared with the classical [illegible] procedure δ^M. The comparisons of the performance of δ^B with [illegible] Monte Carlo studies, indicate that the procedure [illegible] efficiency and smaller expected size of the selected subset. [illegible] indicate that δ^B is robust when the true distributions are not normal [illegible] some other symmetric distributions such as, the logistic, the double exponential (Laplace) and the gross error model (the contaminated distribution) [illegible]

-- 1 OF 1
-- 11 - TITLE: (U) MULTIPLE DECISION THEORY, ORDER STATISTICS AND RELATED
-- PROBLEMS
-- 1 - AGENCY ACCESSION NO: DN423556
-- 5 - SECURITY OF WORK: UNCLASSIFIED
-- 12 - S + T AREAS:
-- 008700 MATHEMATICS AND STATISTICS
-- 011700 OPERATIONS RESEARCH
-- 21E - MILITARY/CIVILIAN APPLICATIONS: CIVILIAN
-- 2 - DATE OF SUMMARY: 06 JAN 81
-- 39 - PROCESSING DATE (RANGE): 14 JAN 81
--10A1 - PRIMARY PROGRAM ELEMENT: 61153N
--10A2 - PRIMARY PROJECT NUMBER: RR01405
--10A2A - PRIMARY PROJECT AGENCY AND PROGRAM: RR01405
--10A3 - PRIMARY TASK AREA: RR0140501
--10A4 - WORK UNIT NUMBER: NR-042-243
--17A1 - CONTRACT/GRANT EFFECTIVE DATE: JUL 75
--17A2 - CONTRACT/GRANT EXPIRATION DATE: JAN 84
-- 17B - CONTRACT/GRANT NUMBER: N00014-75-C-0455
-- 17C - CONTRACT TYPE: COST TYPE
--17D2 - CONTRACT/GRANT AMOUNT: $ 53,166
-- 17E - KIND OF AWARD: EXT
-- 17F - CONTRACT/GRANT CUMULATIVE DOLLAR TOTAL: $ [illegible]

-- 19A - DOD ORGANIZATION: OFFICE OF NAVAL RESEARCH (436)
-- 19B - DOD ORG. ADDRESS: ARLINGTON, VA 22217
-- 19C - RESPONSIBLE INDIVIDUAL: WEGMAN, E J
-- 19D - RESPONSIBLE INDIVIDUAL PHONE: 202-692-4315
-- 19U - DOD ORGANIZATION LOCATION CODE: 5110
-- 19S - DOD ORGANIZATION SORT CODE: 35032
-- 19T - DOD ORGANIZATION CODE: 265570
-- 20A - PERFORMING ORGANIZATION: PURDUE UNIVERSITY DEPT OF STATISTICS
-- 20B - PERFORMING ORG. ADDRESS: LAFAYETTE, IN 47907
-- 20C - PRINCIPAL INVESTIGATOR: GUPTA, S S
-- 20D - PRINCIPAL INVESTIGATOR PHONE: 317-494-[illegible]
-- 20U - PERFORMING ORGANIZATION LOCATION CODE: 1802
-- 20H - PERF. ORGANIZATION TYPE CODE: 6
-- 20S - PERFORMING ORG. SORT CODE: 50418
-- 20T - PERFORMING ORGANIZATION CODE: 231750
-- 22 - KEYWORDS: (U) TWO-STAGE PROCEDURES (U) MULTIPLE DECISION (U)
-- RANKING AND SELECTION (U) ORDER STATISTICS (U) RELIABILITY
-- 37 - DESCRIPTORS: (U) DECISION MAKING (U) DECISION THEORY (U)
-- DISTRIBUTION THEORY (U) GOVERNMENT PROCUREMENT (U) LOGISTICS
-- (U) OPERATIONS RESEARCH (U) PROBABILITY (U) QUALITY CONTROL
-- (U) RELIABILITY (U) SAMPLING (U) STATISTICS (U) STATISTICAL
-- ANALYSIS
--......

APP[illegible]
DI[illegible]

DATE

FILMED